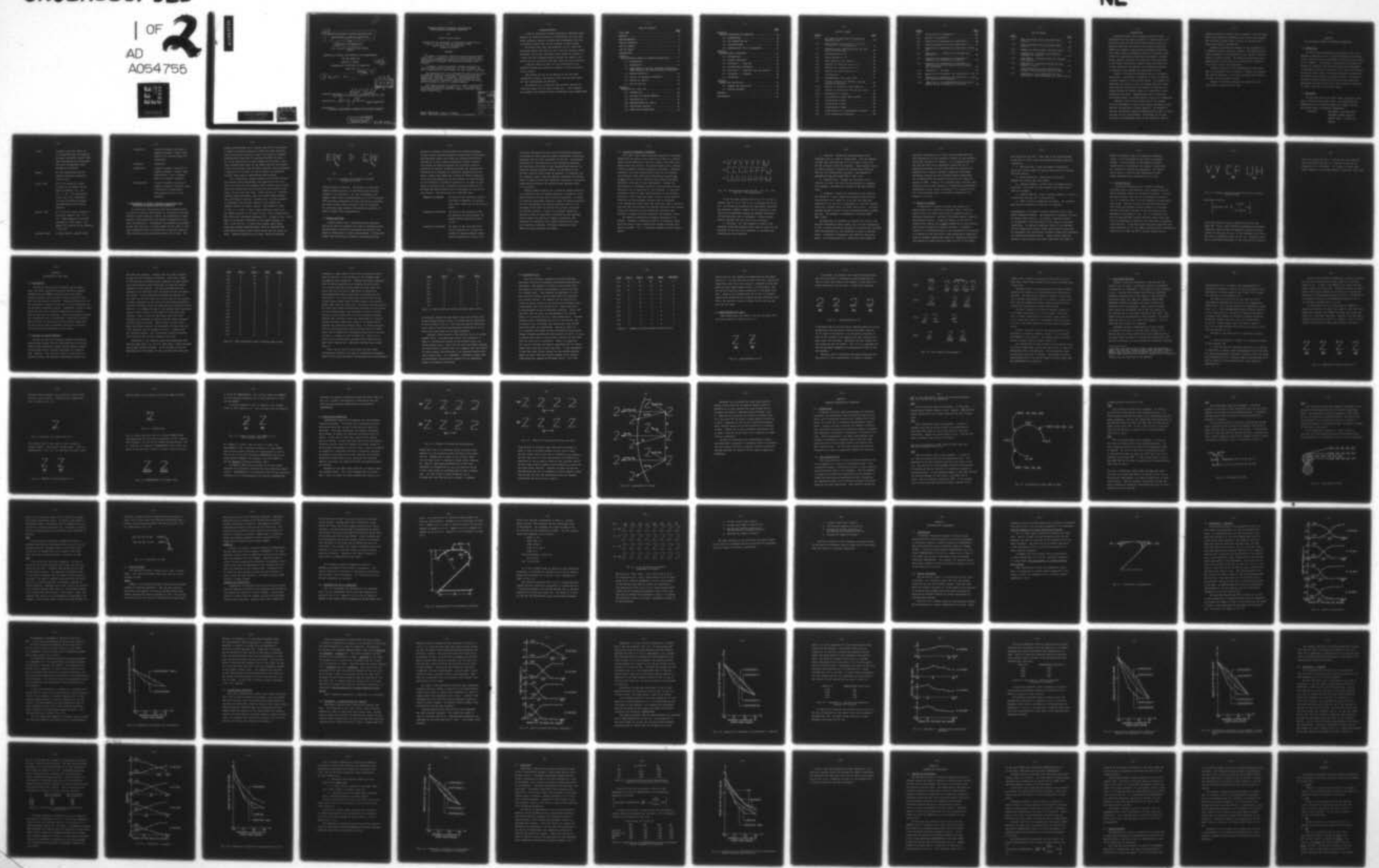
AD-A054 755 MASSACHUSETTS INST OF TECH CAMBRIDGE DEPT OF ELECTRI--ETC F/G 6/4
SIMULATION METHOD OF FEATURE SELECTION FOR UNCONSTRAINED HANDPR--ETC(U)
MAY 77 R T BABCOCK

UNCLASSIFIED NL

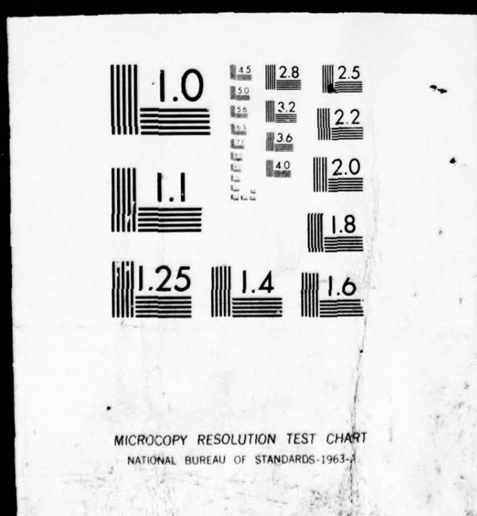
1 OF 2
AD
A054755



IFIED

1 OF 2

AD A054755



DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

0

6

SIMULATION METHOD OF FEATURE SELECTION FOR UNCONSTRAINED HANDPRINTED CHARACTERS.

by

10 Robert Tornow/Babcock

B.S., United States Military Academy
(1970)

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1977

12 99p.

11 May 77

9 Master's thesis,

Signature of Author..........Robert T Babcock..........
Department of Electrical Engineering, May 12, 1977

Certified by..
Thesis Supervisor

Accepted by...
Chairman, Departmental Committee on Graduate Students

409 490

SIMULATION METHOD OF FEATURE SELECTION FOR UNCONSTRAINED HANDPRINTED CHARACTERS

by

Robert Tornow Babcock

Submitted to the Department of Electrical Engineering on May 12, 1977 in partial fullfillment of the requirements for the Degree of Master of Science

ABSTRACT

A theory of character recognition and associated methodology which is expected to lead to a machine algorithm for recognizing unconstrained handprinted characters is reviewed and expanded to encompass the previously excluded Arabic numerals.

A computer character generator package developed for this and future work is described. By way of example, its use in generating the test stimuli for this work is explained.

The character pair 2-Z is selected for study in this work. A systematic method of describing the essential difference between the characters is employed. Three psychophysical experiments, labeling, reaction time and goodness are described and used to determine a quantitative rule for distinguishing between 2s and Zs in neutral context.

This thesis serves as a guide for initial investigation of other intercharacter boundary rules. The collection of rules should eventually describe a general character recognition algorithm.

THESIS SUPERVISOR: Barry A. Blesser
TITLE: Associate Professor of Electrical Engineering

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to Professor Barry Blesser, Dr. Robert Shillman and Ted Kuklinski for their continual patience, support, guidance and encouragement and also to Professor Murray Eden for his interest in this subject.

My loving wife, Jane, who deserved a lot of credit for being both mother and father to our 2 1/2 year old sons, Brian and Scott, during many long days, never seemed to receive it. I thank her most sincerely because without her support, completion of this work would have been infinitely more difficult.

A special thanks is given to Lois Iannaccone for her beautiful editing and typing and calmness in making revisions to this thesis.

Many thanks are due to the members of RLE for their enthusiastic support, particularly John Cook and Andy Poynor for their photography and drafting.

My presence as a student was made possible through the financial support of the United States Army. This research was funded by the National Science Foundation, Grant ENG74-24344.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Although machines which recognize machine printed characters have been quite successful in recent years, with high input rates and low error rates, there are still no machines which can come close to the accuracy with which humans can recognize unconstrained handprinted characters. This thesis reviews a developing theory of character recognition for unconstrained handprinted characters, includes Arabic numerals which have not previously been considered in the theory, describes a computer system for generation of test characters and explores the recognition boundary in character space between the troublesome 2-Z pair.

Chapter 1 reviews the developing theory of character recognition. Terminology is expanded to cover the addition of Arabic numerals to the previous theory. No argument for developing the theory in terms of feature detection rather than template matching and a description of the basic difference between the present theory and statistical pattern recognition are given. Methodologies used in developing the present theory and significance of context are summarized.

Chapter 2 gives a brief description of the present state of development of the present theory and makes an argument for the addition of Arabic numerals. An estimate of the most troublesome character-numeral pairs is used to select the pair 2-Z for investigation. An analysis of the ways 2s and Zs are handprinted leads to the conclusion that the

essential difference between them is whether or not the upper half of the character functions as one or two line segments. The difference between plain 2s and Zs is decided upon as being the richest in information.

Chapter 3 describes the results of the development of a computer character generator package. The package simplifies and speeds up the production of test stimuli, formerly done on a drafting board. The developed instruction set is demonstrated in generating the two dimensional trajectory of 2s and Zs used in the experiments.

Chapter 4 describes three psychophysical experiments used to investigate the 2-Z boundary. Results are compared and used to determine a rule which describes that boundary in terms of the parameters varied.

Chapter 5 provides a summary and conclusions and a discussion of possible future work.

CHAPTER 1

THE DEVELOPING THEORY OF CHARACTER RECOGNITION

1.1 INTRODUCTION

In 1973 Blesser and Shillman [7] reported their use of ambiguous characters to aid in the determination of the functional (identity bearing) attributes of the 26 capital letters of our alphabet. The concept of determining the presence or absence of functional attributes from physical measurements of ambiguous characters is the basis for a theory of character recognition which should, when fully developed, describe how to achieve human accuracy in recognizing hand - and machine - printed characters without having to train the population generating the input to constrain their characters according to a more rigid set of rules than normal.

1.2 TERMINOLOGY

Shillman [8] defined the terms letter, character, allograph, letterform and letter label. His definitions are ammended here to include the addition of numerals to the developing theory of character recognition:

Character	Any graphic sign that may be assigned a letter label or numeral label. A letter or numeral.

Letter	A general term that refers to all letterforms that are commonly assigned a particular letter label. (Some letterforms may be assigned more than one character label depending on context).
Numeral	All the numeralforms that are commonly assigned a particular numeral label.
Letter label	One of the twenty-six labels assigned to the members of the English alphabet; e.g.. "A", "B",..., "Z". The appropriate letter label for a letterform can only be determined through experimentation with literate subjects.
Numeral label	One of the ten labels assigned to the arabic numerals "0", "1",..., "9". Roman numerals are not yet considered in this theory. The number "10" consists of two numerals, "1" and "0".
Character label	A letter label or numeral label.

Letterform	A specific graphic sign that is commonly assigned a letter label. There are an infinite number of letterforms.
Allograph	Letterform.
Numeralform	A specific graphic sign that is commonly assigned a numeral label. (Numeralform is not a standard word of the English language).
Characterform	A specific graphic sign that is commonly assigned a character label. A letterform or numeralform. (Characterform, also, is non-standard).

1.3 SHORTCOMINGS OF OPTICAL CHARACTER RECOGNITION (OCR) MACHINES FOR READING VARIABLE CHARACTERS

In recent years OCR machines have been developed which can achieve virtually 100% accuracy in recognition of machine printed or typed material; however, the performance of these machines drops drastically with moderate degradation of the input characterforms. Furthermore, the machines are designed to work well with only a limited number of type fonts or with input characters which are constrained in physical construction according to some relatively strict set of rules. The

scheme these machines use is usually some sort of correlation or template matching technique in which the input characterforms are labeled according to one of the ideal or archetype characterforms from which it physically differs the least. Due to the sensitivity of these schemes to variations in the input characterforms which do not affect identity and insensitivity to variations which do affect identity, their performance is generally far inferior to that of humans in recognizing variable input such as unconstrained hand printing.

It is conceivable, through analysis of all kinds of hand and machine printed characters, each quantized to an m by n matrix of black and white squares, that a device could be made that would respond with the statistically most probable character, if *any*, *given any* input. Such a device, an idealized template matching machine, could not achieve human accuracy over variations of context. Blesser et al. [1] gave examples of characters that had one identity in one context and another identity in another context. One example is shown in Fig. 1.1. In,Fig. 1.1(a) a human would most likely call the middle character as "D" while in Fig. 1.1(c) a human would most likely call the middle character "P". As seen in Fig. 1.1(b) there is no physical difference in the middle character of Figs. 1.1(a) and (c). In such cases, humans would almost always make correct identifications, while an idealized template matching machine would always assign the same character label. Template matching can, at best, identify characters

(a) (b) (c)

Fig. 1.1 An Example of the Effect of Context on Character Label.

without regards to context. Furthermore, an idealized template matching device, as described, would be quite awesome in its complexity. In a 24 by 24 matrix, for example, there would be approximately 2.5×10^{173} (actually 2^{576}) possible combinations of black and white to consider. For these reasons feature detection must be considered for achieving human accuracy with machine recognition of highly variable input characterforms.

1.4 FEATURE DETECTION

Literate humans have a tremendously well developed ability to read and recognize all kinds of isolated printed characterforms in spite of considerable variation in the physical construction of those characterforms. Shillman [8] argues that the ability of humans to consistently group

physically different characterforms into classes according to their labels without necessarily having seen the particular characterforms before and without the characterforms being necessarily in their normal orientation, refutes a theory that humans use template matching (in the strict sense of measuring the physical differences between the observed characterform and a template) for character recognition and supports a theory that humans use a system of feature detection. Since the exact meaning of the term feature has been obscured in the literature of pattern recognition it is useful to talk about an alternate term, attribute, and three subsets of it. Shillman [8] defines three kinds of attributes in lieu of the term feature. The definitions are summarized here:

Physical attributes	The parts of the character, usually described in geometric or topological terms, which make up the physical image.
Perceptual attributes	The parts of the character which are perceived as being present (or absent) whether physically there or not.
Functional attributes	The parts of the character which act (or function) as if they were present (or absent) whether perceptually or physically there or not.

Functional attributes are the so-called distinctive features of characters; those properties which distinguish a particular character from all other characters in the set. The physical attributes of a characterform do not necessarily determine the functional attributes of that characterform, otherwise an idealized template matching machine (as described in the previous section) would always be accurate. As Kuklinski [3] has rigorously shown, both the physical attributes and various kinds of context affect what functional attributes a particular characterform possesses and therefore what character label it is assigned.

Whether humans actually use a system of feature detection or an adaptable type of template matching is academic; if two models produce the same result they are equivalent. A theory of character recognition based on a model of template matching has the disadvantage that template matching theory is not developed to the point of dealing with variable characters and variable context. A theory of character recognition defined in terms of functional attributes has the advantage that it models the ability of humans to generalize the properties of characters and to modify the decision rules as a function of context. For these reasons the latter theory is being developed and tested.

1.5 THE USE OF AMBIGUOUS CHARACTERS

While most other feature detection theories of character recognition are based on the clustering of data in a feature space and the formation of decision surfaces from some parameters about the clusters, the present theory uses a slightly different approach based on ambiguous characters. Ambiguous characters are those characterforms which can be assigned two character labels with equal probability (actually two or more but for initial simplicity only pairwise ambiguous characters have thus far been considered). Through the determination of ambiguous characters the boundaries between characters can be found more or less directly. If two characterforms, which are good representations of two different characters, differ from each other in only a single physical dimension and the ambiguous character can be determined at an intermediate point along that dimension, then a relationship has been found between the physical attribute (the distance along that dimension) and the functional attribute (the presence or absence of which determines the character label).

The method of determining ambiguous characters is as follows: Generate a trajectory (series) of characters along which the label of the characterforms changes from one character to another. Fig. 1.2 contains examples of three trajectories.

Fig. 1.2 Characterforms along the "V" - "Y", "C" - "F" and "U" - "H" trajectories.

In (a) the label changes from "V" to "Y", in (b) it changes from "C" to "F" and in (c) it changes from "U" to "H". By performing psychophysical experiments, determine the character, whether a member of the trajectory or one that could be generated between two adjacent characters of the trajectory, that is ambiguous. Various psychophysical methods that are useful are as follows:

1. Pointing. Tell the subjects to point to the character along the trajectory that looks as much like the character at one end as the character at the other end. Calculate the mean character.

2. Labeling. Present the characterforms of the trajectory one at a time in random order. Tell the subject to label each characterform as A or B (where A and B are the two character labels of interest). After several trials or several subjects, estimate the probabilities P(A) and P(B) for each characterform and plot. The boundary is estimated at the point where P(A) = P(B) = 0.5.

3. Reaction time. Present the characters as for labeling and record the time from presentation to response. The boundary is estimated at the peak of the mean reaction times.

4. Goodness. Present the characters as for labeling but instead of a label, obtain the subjects rating of each characterform, on a scale, as to how well the character represents A and how well it represents B. Half the subjects should do A ratings first, the other half B first. Calculate the mean goodness ratings of each characterform $\bar{G}(A)$ and $\bar{G}(B)$ and plot. The boundary is estimated at the point where $\bar{G}(A) = \bar{G}(B)$.

By describing the boundary in terms of the value of the physical parameter(s) being varied, (ℓ_1/L for the trajectories in Fig. 1.2) the presence or absence of the functional attribute which distinguishes the two characters is found in physical terms. A physical to functional rule (PFR) has been determined. It is postulated that a relatively small number of

functional attributes can be found (see Shillman [8]) such that combinations of the presence or absence of each uniquely define the characters of interest. It is expected that all the relevant PFRs can be determined through the use of ambiguous characters and that those PFRs can eventually be combined into a general character recognition algorithms. Continuing research has been directed towards determining and validating PFRs and towards determining the effects of context on PFRs. Finally, it is expected that a machine that achieves close to human accuracy in recognizing unconstrained handprinted is most likely to succeed if it is based on this (or a very similar) theory of character recognition (see Blesser et al, [2]).

1.6 EFFECTS OF CONTEXT

As mentioned in previous sections the context of a characterform has a bearing on character label. There are several levels of context which play an important role in character recognition. Kuklinski [3] differentiates linguistic and graphical context. Linguistic context is any form of context which affects the _a priori_ probability of any character label being assigned to an unseen character. Linguistic context is a significant factor in the recognition of characters as evidenced by the fact that a large percentage of characters could be randomly replaced by blanks in a paragraph of English text before the reader would be unable to fill in the blanks

and reconstruct the text. Clues used in the reconstruction of the text in such a case are obtained from many levels of linguistics:

1. Word spelling. Given the number of characters, there are a finite number of English words which will satisfy the blank spaces.

2. Word probability. Some words are much more probable in ordinary text than others.

3. Sentence grammar. Certain rules of grammar must be satisfied. Some forms are more probable than others while some are forbidden.

4. Sentence meaning. Even if syntax (grammar) is correct, a particular word may not fit semantically.

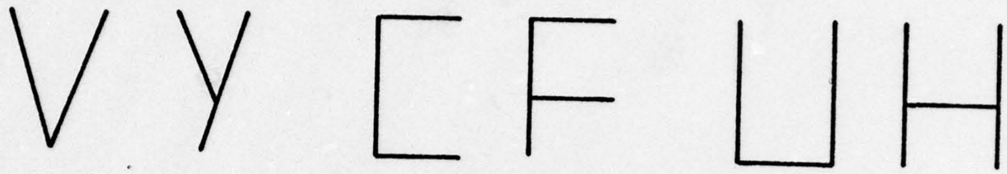
5. Higher levels of semantics and syntax. Two sentences make sense by themselves but may not fit together.

Incorporating such levels of context awareness into a character recognition machine may eventually be necessary in order to achieve human accuracy. At present, linguistic context is disregarded in the development of this theory of character recognition. In contrast, graphical context is an important form of context which must be taken into account to guarantee an improvement over conventional character recognition techniques for unconstrained hand-printing. Simply stated, graphical context is the printing style which conditions the reader to

expect a particular style for subsequent characters. Figure 1.1 was an example of the effect of graphical context. Initially PFRs are determined in neutral context and as a refinement, the changes in those PFRs as a function of context are determined. The latter can be used as a test for the commonality of PFRs across character pairs (see Blesser et al. [2]).

1.7 THE PFR FOR LEG

Shillman [8] proposed a set of twelve functional attributes which, with modifiers, could uniquely describe the 26 uppercase letters of the English alphabet. No claim was made, however, that the set of attributes is unique, i.e. it is possible that it is not complete or, alternately, not minimal. The set, to date, has not been modified but still needs considerable scrutiny before it is verified. The only functional attribute of the set which has been thoroughly investigated is LEG. Fig. 1.3 contains three pair of characterforms (called confusion pairs) which seem to be distinguished by whether or not they have a descending line extension. It has been shown [2,3,4,8,9] that the functional attribute LEG is the common attribute which distinguishes these pairs and that the PFR in neutral context can be



(a) (b) (c)

Fig. 1.3 Confusion pairs distinguished by the functional attribute LEG.

expressed as follows:

$$\left[\text{Functional LEG:} \quad \frac{\ell_1}{L} \underset{\text{Not Present}}{\overset{\text{Present}}{\gtrless}} 0.17 \right]$$

Three tests [2] were used to establish and verify the above PFR. First, three different psychophysical paradigms (labeling, reaction time and goodness) yielded consistent results for the value of ℓ_1/L at the crossover point within each pair of characters shown in Fig. 1.3. Secondly, there was no significant difference in the value of ℓ_1/L crossover

value for each of the pairs. Thirdly, the ℓ_1/L crossover value varied for each pair as a function of context by essentially the same amount. The success of these tests lends credence to the present theory of character recognition.

CHAPTER 2

SELECTION OF A TEST CASE

2.1 INTRODUCTION

The PFR for LEG has been determined and validated. Other functional attributes have been partially explored (CLOSURE [8] and SYMMETRY [10]) but until the present research, numerals were excluded as were variations along more than one physical dimension. Strong evidence exists for the inclusion of numerals at this point in the development of the character recognition theory. Furthermore, strong evidence exists for the need to consider trajectories of characters in more than one dimension. This and subsequent chapters are intended not only to investigate a particular character pair, but also to include much of the insight gained thus far as an aid to other researchers desiring to contribute to the continuing development of this theory.

2.2 THE NEED TO INCLUDE NUMERALS

Neisser and Weene [5] collected samples of uppercase unconstrained (except that the characters had to be mostly within a box) handprinted letters and the Arabic numerals and presented them as isolated characters to a group of nine subjects. They put their results in the form of a confusion matrix, one dimension representing stimuli and

the other the response. Elements were the total responses by type for each category of stimulus. There were a total of 647 characters presented with 9 responses to each (except that some subjects became inattentive at times and made no response to some stimuli). The group averaged about 96% accuracy on an individual basis and about 97% with the pooled best guess. It was found that discrimination between "I" and "one" and "O" and "zero" was so poor that for the purposes of their experiment the two pairs were considered equivalent; incorrect responses within the pairs were not counted as errors. Table 2.1 is a summary by type of the eighteen most significant errors, listed in order of the number of times each error occurred. Type 1 errors are the errors where the first character was presented as stimulus and the second character was the response in error. Type 2 errors are the errors where the second character was presented as stimulus and the first character was the response in error. A modal error was recorded whenever, for a particular stimulus presented, more incorrect than correct responses were given.

Two-thirds of the confusion pairs were associated with only one type error, e.g. for V-Y, seven type 1 errors occurred and no type 2 errors occurred. In all but two of the just mentioned pairs the number of errors was from five thru nine.

PAIR	TYPE 1	TYPE 2	TOTAL	MODAL
Z-2	0	29	29	3
U-V	6	12	18	1
J-U	1	10	11	1
X-Y	3	8	11	1
Y-4	11	0	11	1
Q-2	9	0	9	1
H-N	0	9	9	1
G-6	3	5	8	1
S-5	2	5	7	1
T-7	7	0	7	1
V-Y	7	0	7	1
4-9	7	0	7	1
B-R	6	0	6	1
E-Y	6	0	6	1
C-F	5	0	5	1
J-V	0	5	5	1
V-X	4	0	4	1
C-L	4	0	4	1

Table 2.1 Most significant errors including modal errors.

Therefore, a large number of the errors occurred at times when the stimuli, in the consensus of the subjects, best represented other characters. Although the authors reported that no script characters were presented to the subjects it is likely that the Q which caused the modal error was really one form of script Q ($\mathcal{Q}$) which is essentially indistinguishable from a 2 and that it accounted for all nine errors. It is also likely that the left side of the U that caused a modal error in being called J, was significantly lower than the right side. It appears that the majority of errors were caused by characters which were so poorly made by the originator that, when presented in neutral context, they were significantly over the inter-character boundary. Assuming that the modal errors are unavoidable in a neutral context character recognition scheme it is interesting to consider the remaining confusion pairs. By further assuming that the maximum possible number of errors were also modal errors and by temporarily disregarding those errors, the data of Table 2.1 reduces to that of Table 2.2. Table 2.2 is surely an over-conservative estimate of the number of non-modal errors however two important inferences may be made from it.

First, the U-V pair is very likely the most often misrecognized character pair near an intercharacter boundary (modal errors being considered significantly over the boundary).

PAIR	TYPE 1	TYPE 2	TOTAL
Z-2	0	2	2
U-V	6	3	9
J-U	1	1	2
X-Y	3	0	3
Y-4	2	0	2
G-6	3	0	3
S-5	2	0	2
C-L	4	0	4

Table 2.2 Most significant errors excluding modal errors.

A successful attempt has been made by Suen and Shillman [12] using weighted feature vectors to perform machine recognition of thick stroke Us and Vs. They reported machine recognition error rates of digitized unconstrained handprinted Us and Vs lower than human rates for the same characters.

Secondly, fifty percent of the confusion pairs are letter-numeral pairs. Even when modal errors are included as in Table 2.1 thirty-three percent of the confusion pairs are letter-numeral pairs. Along with the fact that Arabic numerals are used freely in English text, this suggests that the problem of letter-numeral discrimination is an important problem that must be dealt with. It is therefore reasonable at this time to specifically include Arabic numerals in the developing theory of character recognition.

2.3 SELECTION OF 2-Z

There are different approaches one might take when designing a machine to recognize unconstrained hand-printed characters. One approach is collect a lot of data about the shapes of characters, and attempt to cluster the data in a feature space. This approach is problematic in that more and more features must be added until all characters are uniquely defined. For example, all As might be described as having closure at the top and two descending legs, a description that also applies to Rs. These cases must be disambiguated by use of additional features. Rather than save this step for last, the methodology of the present theory is to initially focus on such problem cases. Although confusion matrix techniques may not be a satisfactory way to extract a set of features [2] they do give some insight into which pairs are troublesome. Looking at Table 2.1, it is apparent that humans do have difficulty distinguishing between isolated unconstrained hand printed Zs and 2s; in fact, it appears that there is more trouble with the 2-Z pair than with any other pair of characters. Table 2.3 summarizes the results of Neisser and Weene's experiment for all letter-number pairs that had at least one error. If the assumption is made that modal errors account for the maximim possible number of errors then the G-6 pair appears to be the most troublesome with respect to non-modal errors. It is not

PAIR	TYPE 1	TYPE 2	TOTAL	MODAL	NON-MODAL
Z-2	0	29	29	29	2
Y-4	11	0	11	9	2
Q-2	9	0	9	9	0
G-6	3	5	8	5	3
S-5	2	5	7	5	2
T-7	7	0	7	7	0
S-2	2	0	2	0	2
A-9	0	1	1	0	1
O-6	0	1	1	0	1
P-9	0	1	1	0	1
R-7	0	1	1	0	1

Table 2.3 Summary of letter-numeral confusion pairs.

likely that all nine subjects misrecognized the same three twos as Zs. If they did, the 2-Z pair still has a relatively significant error rate with respect to non-modal errors compared to other letter-numeral pairs. If they didn't, then the 2-Z pair probably assumes prominence in non-modal error rate. Based on the evidence that the 2-Z pair is troublesome and the methodology of studying the most difficult cases first, the decision is made to choose the 2-Z pair as the next case to study.

2.4 CHARACTERFORMS OF 2 AND Z

When handprinted, the letter Z has only one basic form with one variation as shown in Fig. 2.1.

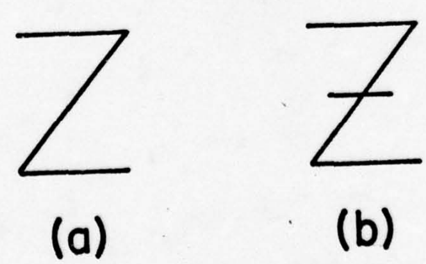


Fig. 2.1 Characterforms of "Z"

In contrast, the numeral 2 has several distinct forms, each of which could be considered a good representation. Fig. 2.2 shows four forms of the numeral 2, each quite different physically from the other. Wright [13] conducted

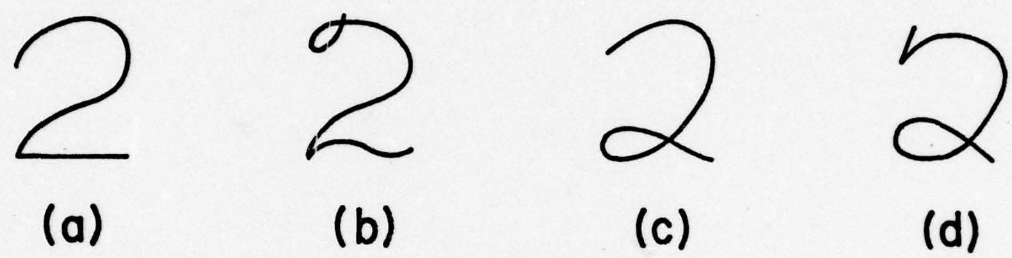


Fig. 2.2 Characterforms of "2"

a thorough study of the way English speaking people of various occupational and educational levels write Arabic numerals. In his study of the numeral 2, he classifies variations in each of four regions of the numeral; the four regions being head, stem, turn and base. Admittedly the four regions are not unique but they are useful. Fig. 2.3 is a condensed version of the major categories of construction of each of the four regions.

The head, which is generally the stroke comprising the top portion of the characterform, is simple or complex;

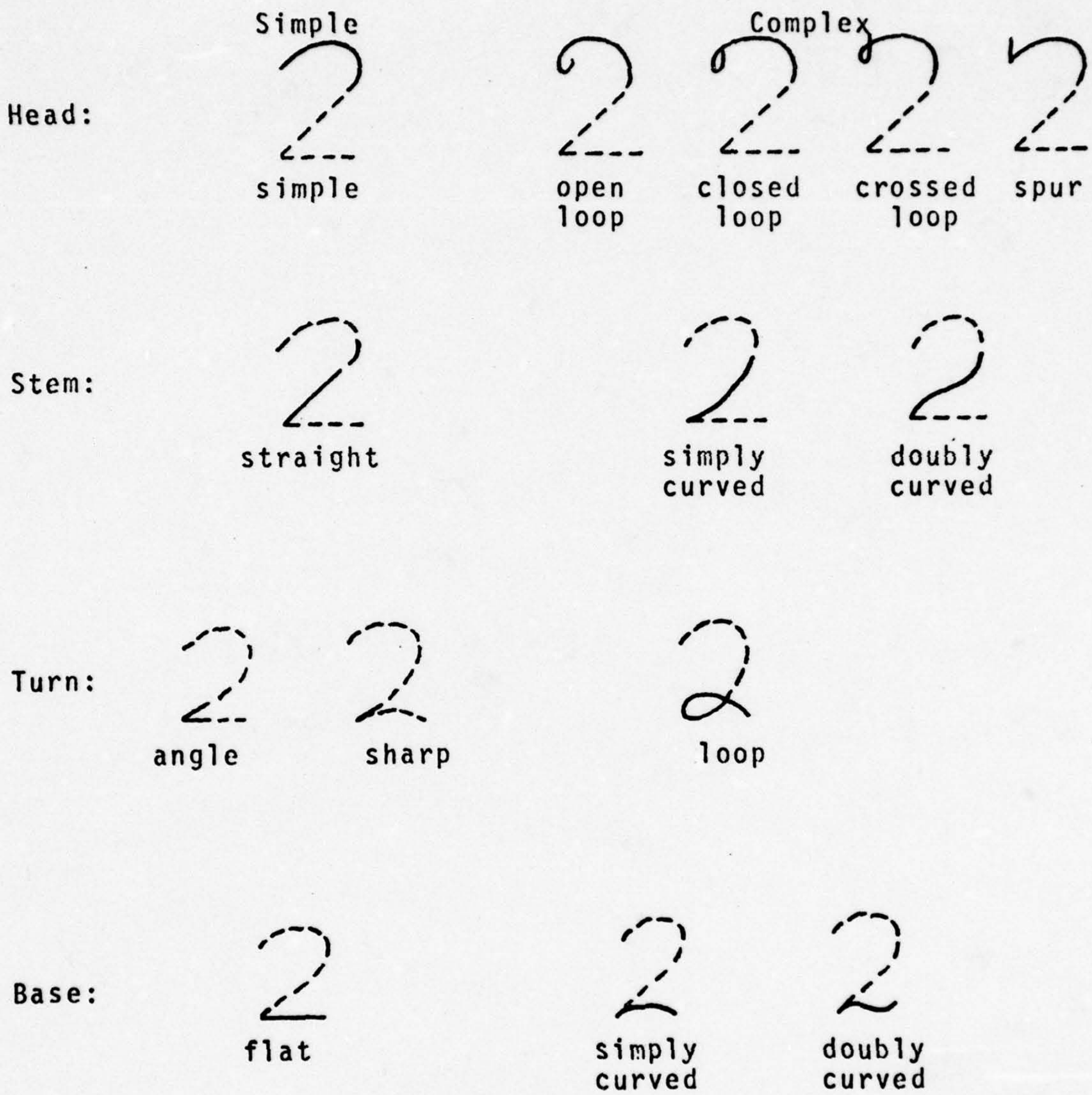


Fig. 2.3 Four regions of the numeral 2

complex when it contains a marker of some kind such as an open loop, closed loop, crossed loop or spur and simple when it does not.

The stem, which is a continuation of the head, extends downward and leftward, not including the turn which begins the base. The stem also is simple or complex; simple when it is straight and complex when it is curved. Complex stems fall into two categories; simply curved and doubly curved (curved and swan-necked as Wright called them).

The turn, likewise, is simple or complex; complex if it is looped and simple if it is not looped. The simple turn is further subclassified as sharp or angle; sharp if it is a very acute angle or a non-looped retrace and angle if it is a less acute angle or a slurred counter-cloclwise curve.

Finally, the base, which is the ending stroke of the character, is simple or complex; simple if it is straight and complex if not straight. The complex bases are simply curved or doubly curved as shown in Fig. 2.3.

The separability of turn and base is not absolute since a 2 with a looped turn, such as the one in Fig. 2.3 does not necessarily have a distinct base. A classification incorporating turn and base is "bottom". The bottom is plain or complicated; plain if it has an "angle" turn and "flat" base and complicated if it does not.

2.5 PRELIMINARY ANALYSIS

Because of the many variations in the way that the numeral 2 can be made, a preliminary analysis is warranted to determine the information - rich regions of the characters 2 and Z. Shillman [8] proposed the following functional attribute representation of the letter Z: the functional attributes SHAFT, LEG, ARM, BAY, CLOSURE, NOTCH, HOOK, CROSSING, SYMMETRY and MARKER are either irrelevant or implied by other designations. WELDs should be absent. There should be two INLETS[1], one at the top, opening to the left and one at the bottom, opening at the right. The INLETS should be concatenated at their shores[2]. The bottom INLET may consist of 1, 2 or more line segments. The top INLET may only consist of 2 segments.

Although the above representation is only claimed to uniquely distinguish Z from the other letters, it is obvious that the latter specification about segmentation of the upper INLET was meant to disallow the characterform 2 (which is obviously a 2) from fitting the functional attribute

---

[1] Shillman's description of the twelve physical attributes, SHAFT, LEG, ARM, BAY, CLOSURE, NOTCH, HOOK, CROSSING, SYMMETRY, MARKER, WELD and INLET which correspond to the twelve identically named functional attributes is summarized in the Appendix.

[2] Shores are also explained in the Appendix.

representation of Z. It appears that segmentation is a functional attribute in itself, in that the change of segmentation of the upper INLET of the Z from 2 to 1 changes the identity of the character to 2.

Two other pairs, U-V and S-5 also seem to be distinguished by segmentation, the former by a segmentation change from 1 to 2, and the latter by a segmentation change from 1 to 3. Since SEGMENTATION directly affects identity, it will tentatively be considered a functional attribute. It appears that another of Shillman's proposed functional attributes, MARKER, may play a role in distinguishing 2s and Zs. Complex heads, turns and bases as shown in Fig. 2.3 are embellishments which serve as markers to indicate the presence of a 2.

Inherent in the definition of a functional attribute are:

1. That its presence or absence is a necessary property of the character and

2. That the physical attribute from which it is derived is sufficient to change a good representation of one character to a good representation of another character, where the phrase good representation of a character means a characterform which will be labeled as that character with a high degree of condifence [2].

Certainly the presence of MARKER is a necessary property of Q and its absence from the lower right is a necessary property of O. A good O with MARKER at the lower right becomes a good Q and vice versa. Therefore, subject to experimental verification, MARKER can be said to be a functional attribute for distinguishing between Q and O. Can MARKER be a functional attribute in distinguishing between 2 and Z? The following argument says it is not.

If the addition of one or more physical MARKERs can change the identity of a good Z to a good 2, with other attributes held constant, then MARKER would meet the second condition of a functional attribute. Consider first the embellishments which make the head of a 2 complex. Fig. 2.4 demonstrates the addition of these embellishments to a good Z.

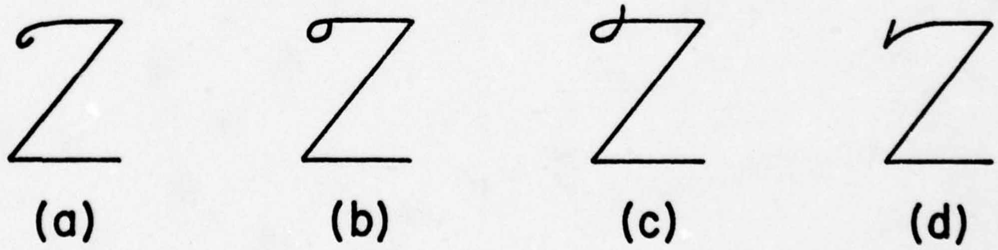


Fig. 2.4 Addition of loops and spurs to Z

Although those characters are not good Zs, none of them could be called a good 2. Consider next a looped turn. This is shown in Fig. 2.5.

Fig. 2.5 Addition of a looped turn to a Z

The characterform is not a good Z, but it cannot be called a good 2. Now consider complex bases. These are demonstrated in Fig. 2.6. As with heads and turns, these

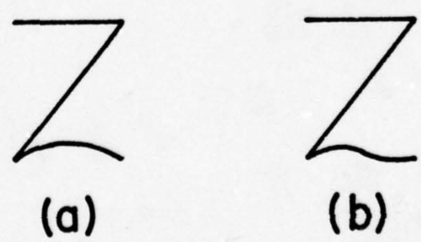


Fig. 2.6 Addition of curved bases to a Z

characterforms are not good Zs but are not good 2s either.

Fig. 2.7 A monstrosity

Fig. 2.7 shows that even with all physical MARKERs added simulataneously the resulting characterform is not a good 2; thus condition 2 is not met for MARKER distinguishing 2 from Z. From Fig. 2.8 removal of SEGMENTATION from a good Z appears to be sufficient to make it a good 2; thus condition

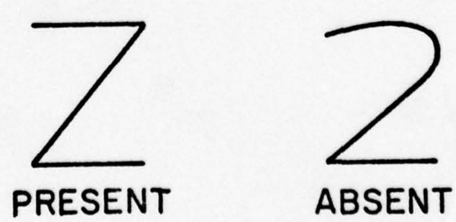


Fig. 2.8 SEGMENTATION of the upper inlet

2 is met for SEGMENTATION. Fig. 2.8 also shows that MARKERS are not necessary properties of 2; thus condition 1 is not met for MARKER.

A similar argument is made in regards to the crossbar which is often added to a Z. Fig. 2.9 shows that the addition

(a) (b)

Fig. 2.9 Further evidence that MARKER is not the functional attribute

of a marker to a good 2 does not make it a good Z and that MARKER is not a necessary property of Z; thus MARKER does not meet either condition for distinguishing Z from 2. This is not to say that functional SEGMENTATION does not include physical MARKER as an argument; only that MARKER is not the essential difference between 2 and Z.

In summary, MARKER has been shown not to meet either condition of a functional attribute for discrimination between 2 and Z. SEGMENTATION is an appropriate name for the functional attribute in 2-Z discrimination since physical SEGMENTATION

describes the physical difference between the good 2 and Z in Fig. 2.8. Further investigation is required to find the rule that maps physical measurements into functional SEGMENTATION

2.6 SIMPLIFYING ASSUMPTIONS

SEGMENTATION should be investigated with the ambiguous character technique. This means that a trajectory must be constructed which has a good 2 at one end and a good Z at the other. Along this trajectory, the physical parameters which appears to affect the functional attribute should be varied. In the case of LEG, only a ratio of two lengths had to be varied. In this case there seem to be more physical variables; various kinds of curvature in various regions of the character and markers of various kinds. Since the presence or absence of a functional attribute is maximally sensitive to geometric variations which affect identity and minimally sensitive to those which do not, the guide for selecting physical attributes are those which have the greatest effect on identity.

Rounding of the upper right corner of a Z seems to have the greatest effect. This can be accomplished in several ways. Fig. 2.10 shows that stem curvature can cause a Z to

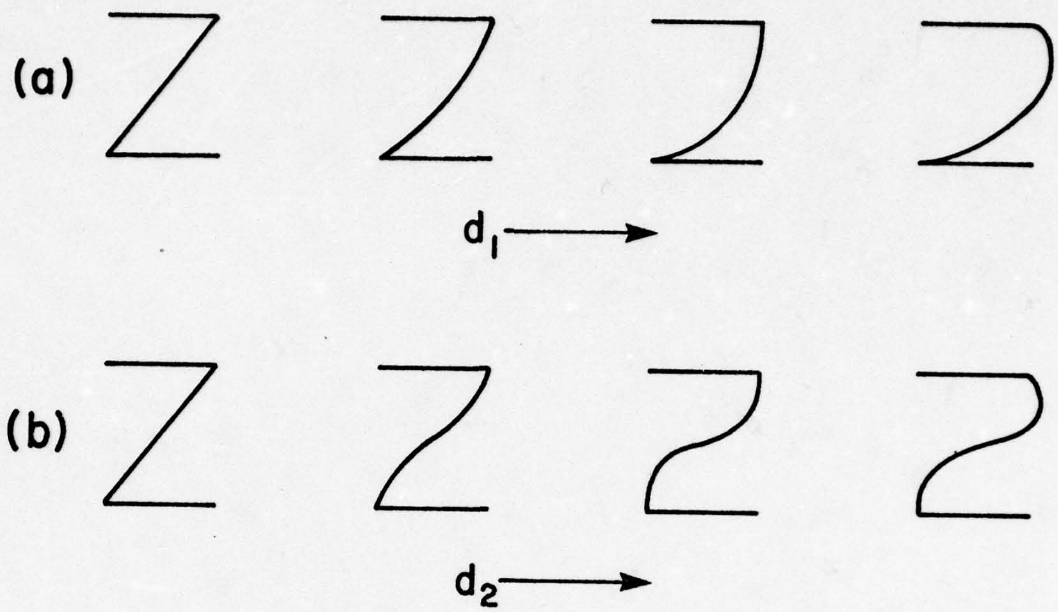


Fig. 2.10 Effects of increasing stem curvature

become more like a 2; increasing single curvature along dimension d_1 and increasing double curvature along dimension d_2. Another way is shown in Fig. 2.11(a). where the top right corner is gradually rounded along dimension d_3. There is some question whether any of the final characters of trajectories d_1, d_2 or d_3 could be called 2 with a high degree of confidence due to their straight tops. For that reason a fourth dimension, d_4 is included such that the top can be rounded. A hypothe-

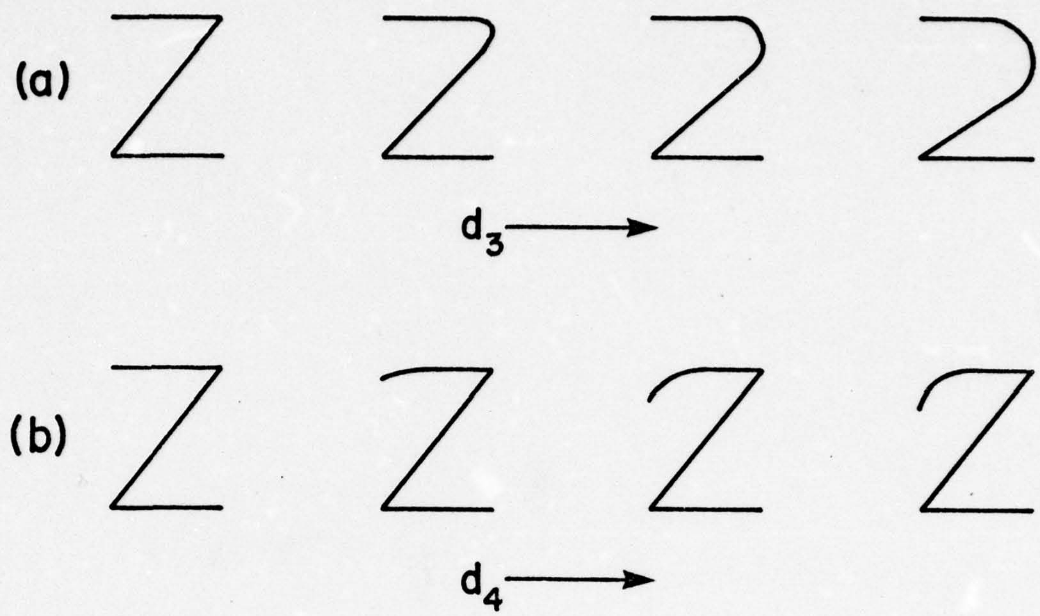


Fig. 2.11 Effects of varying top left and top right.

sized section of character space near the 2-Z boundary is shown in Fig. 2.12 which attempts to show several ideas. The most direct route across the 2-Z boundary is by rounding the top right corner of the characterform via dimension d_3; dimensions d_1 and d_2 require excessive stem curvature before the top right corner is smooth, while dimension d_3 rounds the top right corner directly. Dimension d_4 does not take the characterform over the 2-Z boundary by itself; however, dimension d_3 does not take the characterform over the boundary sufficiently far for it to be a good 2.

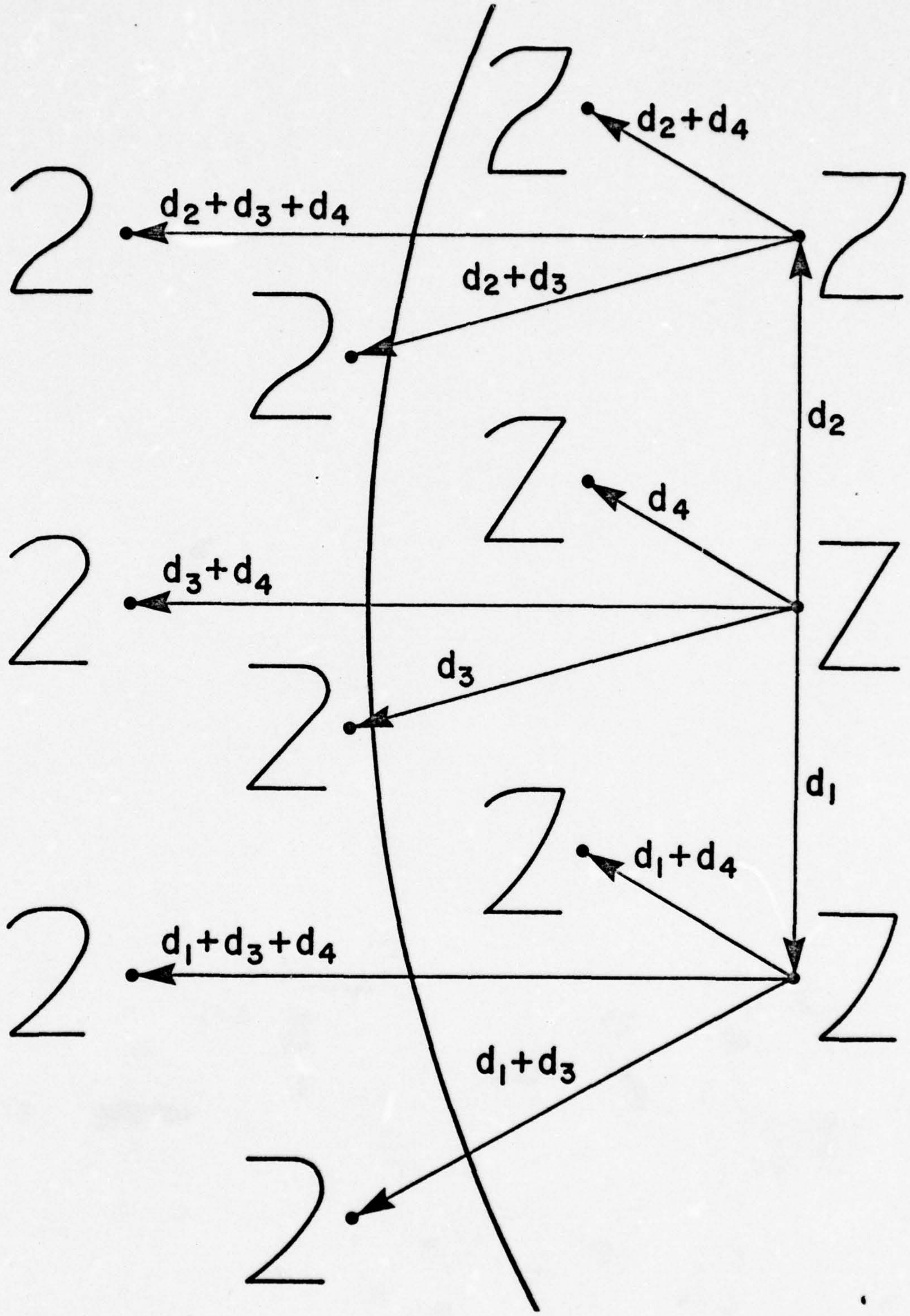


Fig. 2.12 Hypothesized 2-Z space

Dimension d_3 is selected for study since identity change is most sensitive to physical changes along it. Dimension d_4 is also selected for study to meet the requirement of finding a trajectory which includes a good 2. Dimension d_3 and d_4 are not initially combined since it is not clear how to obtain a single measure for the two of them. Dimensions d_1 and d_2 are initially disregarded since identity appears to be less sensitive to variations in them. Effects due to stem variations and markers such as loops, spurs and curved bases may be analyzed and added later as refinements.

In order to investigate the intercharacter border a two-dimensional trajectory of characters should be generated. The next chapter describes a computer character generation package developed for general use and used to create this trajectory.

CHAPTER 3

COMPUTER GENERATION OF CHARACTERS

3.1 INTRODUCTION

A computer character generation package was developed to provide the experimenter with an efficient means to produce a set of characters according to a precise set of parameters for direct use in conducting psychophysical experiments. Features include a powerful and compact instruction set, interactive generation of characters segment by segment, bulk generation of character trajectories in accordance with user - defined parameters, high speed scope display, bulk storage of characters on computer tape in modified chain code and output on the Calcomp digital line plotter with Mars technical pens. The following section describes the instruction set used for generating segments of characters.

3.2 THE INSTRUCTION SET

This instruction set is used to create characters, one or two segments at a time, on a rectangular array. Each instruction assumes a current point which is 0,0 initially. Each instruction takes 2 to 4 arguments. The 1st argument is always the terminating x-coordinate and 2nd argument is always the terminating y-coordinate for that instruction. The terminating point of an instruction becomes the current point for the next instruction. Arcs subtend a maximum of

180^{0} in each instruction. Angles are interpreted modulus 180, e.g. +90 and -90 are equivalent.

JUMP

This instruction takes two arguments. It moves the current point without drawing a line. Example : JUMP 100 100 causes 100,100 to be the current point for the next instruction.

LINE

This instruction takes two arguments. It draws a straight line from the current point to the terminating point and leaves that as the current point for the next instruction. Assume the current point is 0,0. LINE 100 100 draws a straight line from 0,0 to 100,100.

The next 3 instructions, ARCR, ARCO and ARCT, draw arcs and are illustrated in Fig. 3.1.

ARCR

This instruction takes three arguments. It draws an arc from the current point to the terminating point. The magnitude of the third argument specifies the radius of the arc. The sign of the third argument specifies the direction of rotation of the arc; positive is clockwise, negative is counterclockwise. If the specified radius is less than one-half the distance between the two points, a straight line is drawn. The arc subtends a maximum of 180^{0}. If the current point is 50,0 then ARCR 000 050 -50 draws a quarter circle

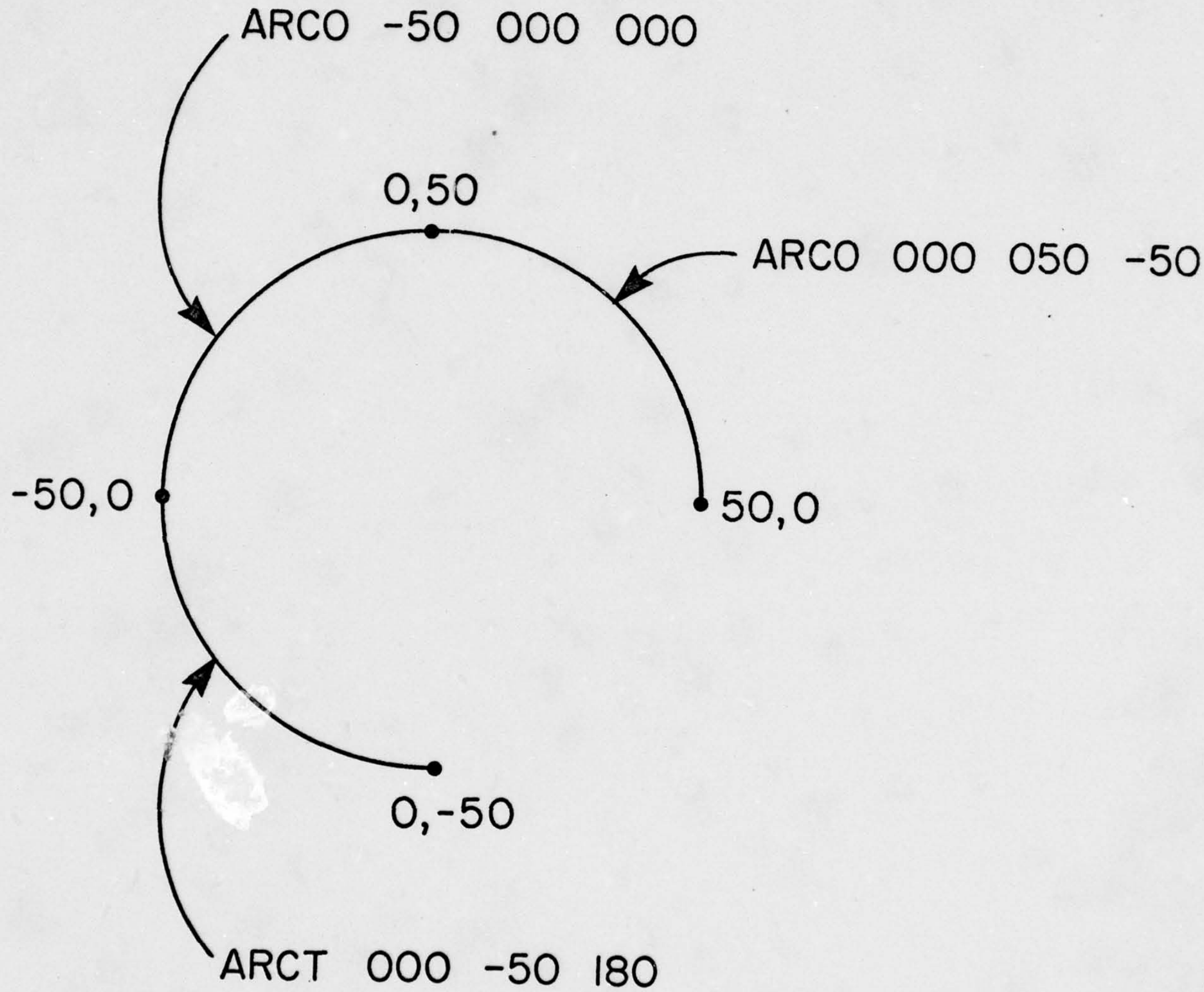


Fig. 3.1 Illustration of ARCR, ARCO and ARCT

counterclockwise from 50,0 to 0,50.

ARCO

This instruction takes three arguments. It draws an arc from the current point to the terminating point. The third argument specifies the angle at the originating (current) point in degrees. The arc will depart the originating point at the specified angle or angle plus 180° such that the arc subtends no more than 180°. If the current point is 0,50 then ARCO -50 000 000 draws a quarter circle counterclockwise from 0,50 to -50,0.

ARCT

This instruction takes three arguments. It draws an arc from the current point to the terminating point. The third argument specifies the angle at the terminating point in degrees. The arc will arrive at the terminating point at the specified angle or angle plus 180° such that the arc subtends no more than 180°. If the current point is -50,0 then ARCT 000 -50 180 draws a quarter circle counterclockwise from -50,0 to 0,-50.

The next 3 instructions, A1A2, R1AN, and R2AN each draw a line and an arc such that the arc is tangent to the line at the point of intersection, i.e. they are continuous in slope and direction. They are powerful instructions in that the user is spared the trouble of calculating the point of intersection of the arc and line.

A1A2

This instruction takes four arguments. The third argument is the angle at the originating point and the fourth argument is the angle at the terminating point. It connects the two points with an arc-line or line-arc combination; whichever is possible.

If the current point is 0,100 then A1A2 050 000 180 090 draws a quarter circle of radius 50 clockwise from 0,100 to 50,50 and from there a straight line to 50,0. If the current point is 0,100 then A1A2 100 050 000 -90 draws a straight line from 1,100 to 50,100 and from there a quarter circle of radius 50 clockwise to 100,50. This is illustrated in Fig. 3.2 .

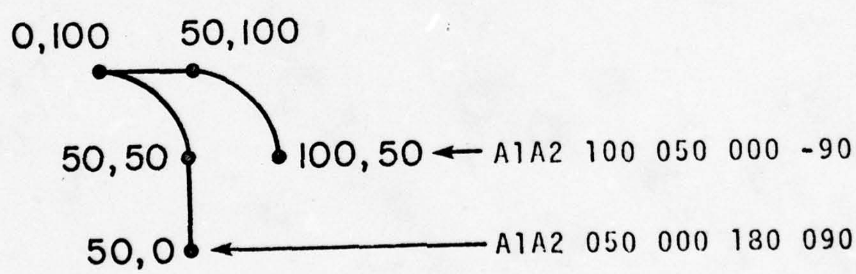


Fig. 3.2 Illustration of A1A2

<u>RIAN</u>

This instruction takes four arguments and draws an arc from the originating point and a line to the terminating point. The magnitude of the third argument is the radius of the arc leaving the originating point. The sign of the third argument is the direction of the arc; positive is clockwise, negative is counterclockwise. The fourth argument is the angle at one point.

If an impossible combination is given it draws a straight line from originating point to terminating point. The instruction is illustrated in Fig. 3.2 for an originating

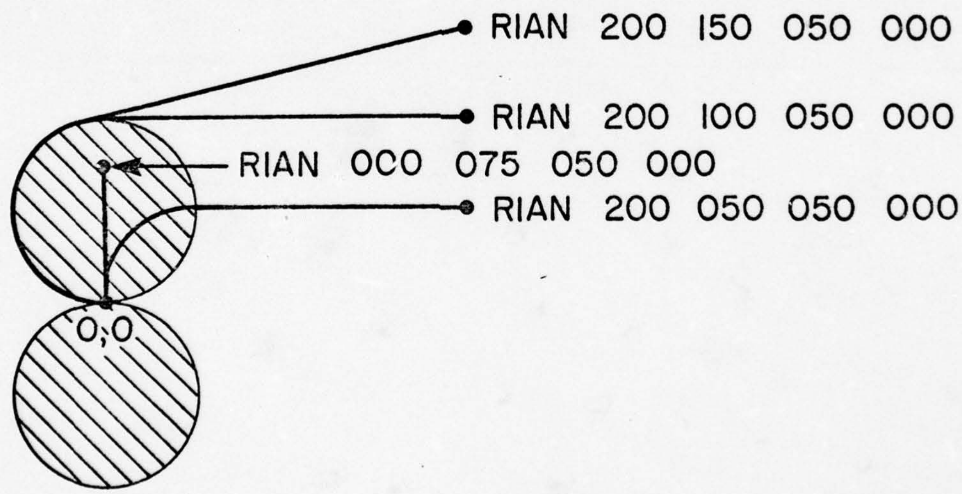


Fig. 3.3 Illustration of RIAN

point of 0,0, an arc of radius +50, an angle of 0 degrees and various terminating points. As shown in the figure, a terminating point in the shaded area is impossible and will result in a straight connecting line. The angle is applied to the originating point or terminating point; whichever is possible.

R2AN

This instruction takes four arguments and draws a straight line from the originating point and an arc to the terminating point. Argument three is the signed radius of the arc and argument four is the angle of one point. Other than those differences R2AN is identical to R1AN.

OVAL

This instruction takes four arguments. It draws a section of an ellipse such that the ellipse axes are horizontal and vertical. Arguments one and two are the terminating point as usual. The magnitude of argument three is the vertical distance from the center of the ellipse to its edge. The sign of argument three determines the direction of the section; positive is clockwise, negative is counterclockwise. Argument four is the ratio of the vertical axis to the horizontal axis times 10. The maximum section of an ellipse that can be drawn is 180 degrees. OVAL, with argument four equal to 10 is equivalent to using ARCR. For example, if the current point is 50,0 then OVAL 000 050 -50

010 draws a quarter circle counterclockwise from 50,0 to 0,50. With current point 50,0 OVAL 000 100-100 020 draws a quarter section counterclockwise from 50,0 to 0,100. This is illustrated in Fig. 3.4.

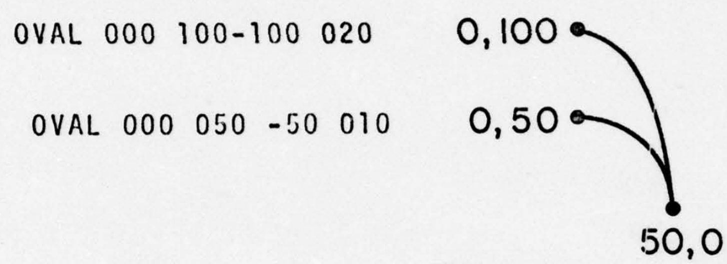


Fig. 3.4 Illustration of OVAL

3.3 SYSTEM FEATURES

The character generator package may be used in several modes. This section describes these modes and the salient features of each.

MANUAL

MANUAL is an interactive graphics mode used for testing schemes of character generation. The user may construct characters, one segment at a time on the MRM storage scope adding, changing and deleting segments at will. At any time the user may list the series of instructions and their argu-

ments which pertain to the displayed character. Displayed characters may be plotted on the Versatec point plotter or Calcomp line plotter as desired. Characters may also be cataloged and stored on disc or tape for future use. Previously stored characters may be recalled from storage for display or plotting as desired. MANUAL mode is entered by making the character generator library available to the loader and by loading the main program MANUAL.

AUTOMATIC

In order to create a series of characters automatically, the user must write a main program in FORTRAN IV. For each character an array must be established containing the instruction and parameters for each segment. Subroutine CALC is called for each segment which interprets the commands, calculates a chain code representation of the character and displays the character on the storage scope. Subroutine FILE is called for each character which puts that character on disc or tape for later plotting. An example program, DOTOP serves as a model program.

VERSATEC OR CALCOMP PLOTS

Trajectories of characters which have been stored under one file name may be plotted on the Versatec point plotter or the Calcomp line plotter in various formats. The Versatec plotter is faster than the Calcomp, but does not make plots with as high of quality. Versatec plots have a definition

of 100 points per linear inch and are made by an electrostatic process. Calcomp plots have a definition of 200 points per linear inch and are made by ballpoint or liquid ink pens. Versatec plotting may be accomplished by making the package library and character file available to the loader by loading the main program PUTVP. Calcomp plotting may be accomplished by making the package library and character file available to the loader and by loading the main program PUTCC. The programs will request information about the character file name, character height and width, desired spacing and margins. Characters will then be plotted in form usable in psychophysical experiments.

This character generator package is written in FORTRAN IV and MACRO 9 for the PDP 9 minicomputer. Sufficient descriptive information is contained on the package tape to make it self-explanatory. The instruction set may be easily ammended and extended.

3.4 GENERATION OF THE 2-Z TRAJECTORY

Fig. 3.5 illustrates the construction of a character that is at an intermediate point along both dimensions d_3 and d_4 of Fig. 2.11. Segment 1-2 is an arc of radius RS (radius of the shore) which is tangent to the horizontal at 2.

Point ? is at the top of the character midway between its left and right extremes. Segment 2-3 is a horizontal straight line. Segment 3-5 is an arc or radius RD (radius of the dock) tangent to segment 2-3 at 3. Segment 5-6 is a straight line tangent to arc 3-5 at 5. Segment 6-7 is a horizontal straight line.

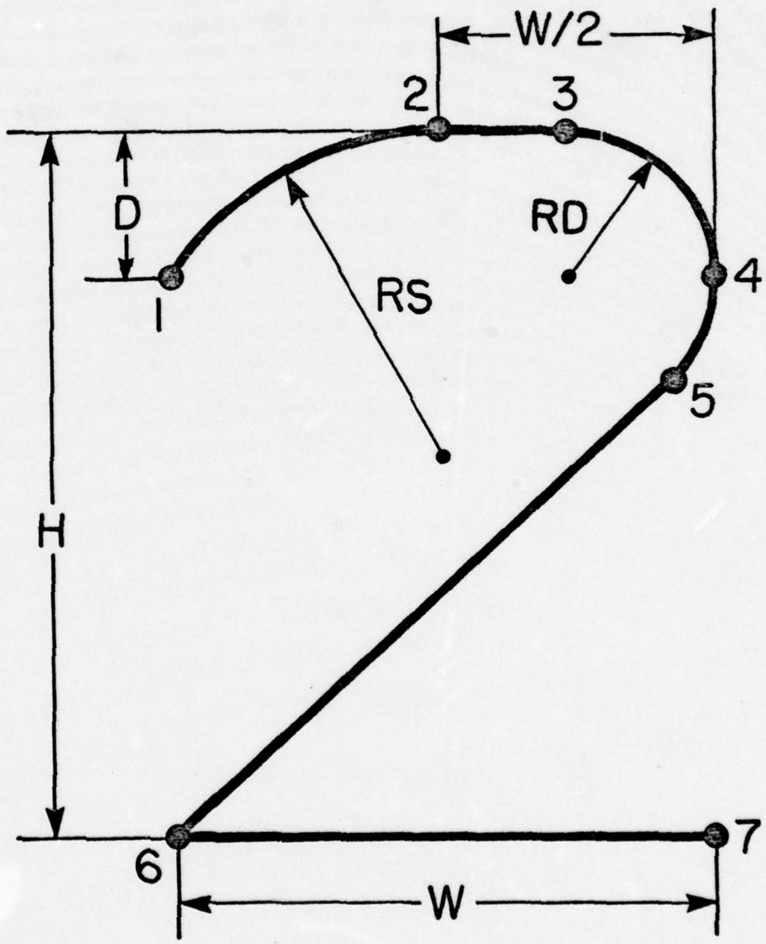


Fig. 3.5 Construction of an intermediate character

Point 4 has the same x-coordinate as point 7. Letting point 6 be 0,0, this character may be constructed using 5 instructions. The arguments are listed in terms of H, W, D and RD, and are separated by commas. The five instructions with arguments are as follows:

JUMP, 0, H-D

ARCT, W/2, H, 0

LINE, W-RD, H

R1AN, 0, 0, RD, 0

LINE, W, 0

This description is valid for

$0 \leq D \leq W/2$

and $0 \leq RD \leq W/2$

It is now a simple matter to generate a two dimensional trajectory of characters by choosing H and W and by picking ranges and increments for D and RD. Such a trajectory is shown in Fig. 3.6.

The original trajectory (Fig. 3.6 is a photographically reduced version) consists of characters 1.3 inches high and 1 inch wide produced on the Calcomp plotter with a .3mm Mars technical pen using black India ink. The values of H and W are 260 and 200 respectively (since the Calcomp increments

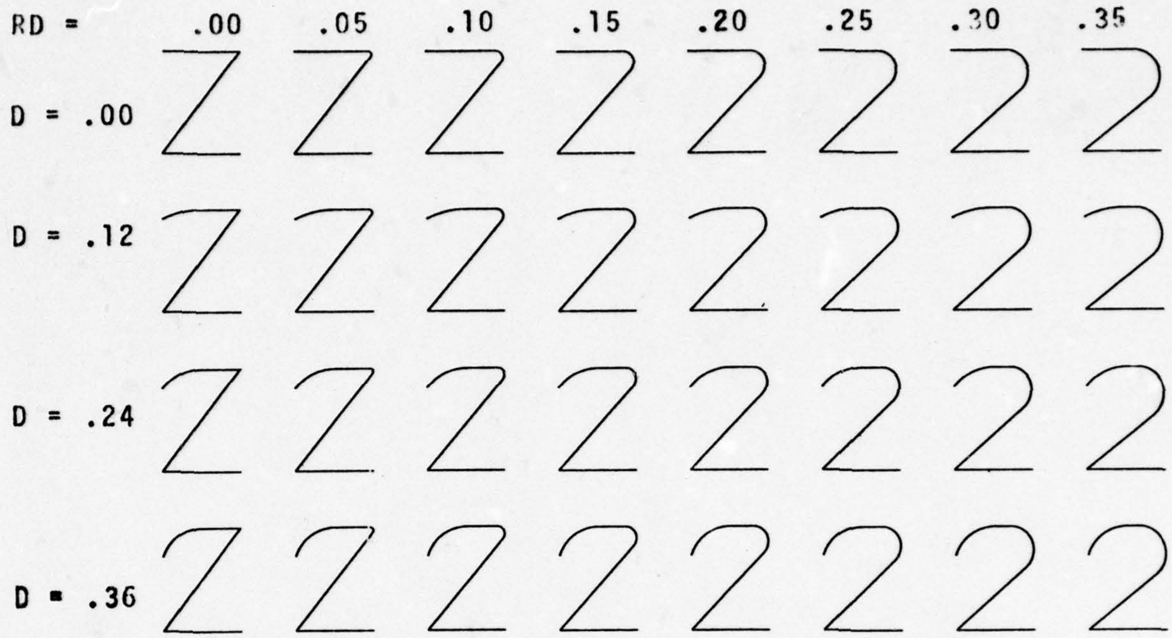


Fig. 3.6 A two dimensional trajectory (parameters in inches)

200 steps per linear inch). D is varied from 0 to 72 in increments of 24. RD is varied from 0 to 70 in increments of 10. Similar characters, 1.3 by 1 inch, plotted on the Versatec were used as stimuli for experiment 1. The characters of Fig. 3.6, 1.3 by 1 inch, plotted individually on the Calcomp and mounted on 4 by 6 inch cards were used as stimuli for experiments 2 and 3. In picking the particular values of D and RD, an attempt was made to do the following:

1. Include a good 2 and a good Z,
2. Include equal numbers of 2s and Zs,
3. Provide for maximum resolution of the boundary as a function of RD and
4. Minimize the number of stimuli.

The final decision on this trajectory was based largely on the advice of colleagues having a great deal of experience with this theory of character recognition.

1. Include a good 2 and a good Z,
2. Include equal numbers of 2s and Zs,
3. Provide for maximum resolution of the boundary as a function of RD and
4. Minimize the number of stimuli.

The final decision on this trajectory was based largely on the advice of colleagues having a great deal of experience with this theory of character recognition.

CHAPTER 4

PSYCHOPHYSICAL EXPERIMENTS

4.1 INTRODUCTION

A character recognition machine will have to make labeling decisions based primarily on physical measurements of characters, however a physical absence of SEGMENTATION does not necessarily mean a functional absence of SEGMENTATION. A rule which maps physical SEGMENTATION into functional SEGMENTATION of 2s and Zs is needed and is a goal of this thesis but is only a preliminary step in examining functional SEGMENTATION. This chapter will describe experiments which attempt to define the boundary between the plain Zs and plain 2s of Fig. 3.6.

4.2 INITIAL HYPOTHESES

The top left of a good Z is straight and its top right is pointed. As it becomes a good 2 the top left becomes curved and the top right becomes rounded. From the standpoint of building a machine, it is desireable to find measures of curvature and roundness which are easily obtained and which can be combined linearly to closely approximate the intercharacter boundary.

From Fig. 3.6 is appears that the intercharacter boundary may be described by a linear combination of D and RD. These

parameters could be quickly measured by a machine on characters like those of Fig. 3.6, therefore hypothesis 1 is formulated; <u>that the boundary is a linear function of D and RD</u>.

Fig. 4.1 shows another measure of roundness and curvature. Area A1 increases as the curvature of the top left increases and area A2 increases as the roundness of the top right increases. The areas are not as easily measureable on the characters of Fig. 3.6, but are applicable to a wider range of characters.

It appears that the boundary might be described by a linear combination of A1 and A2. Therefore hypothesis 2 is formulated; <u>that the boundary is a linear function of A1 and A2.</u>

After formulation of these two hypotheses, experiment 1 is conducted as a pilot study to determine if the range of stimuli is acceptable and to initially compare hypotheses 1 and 2.

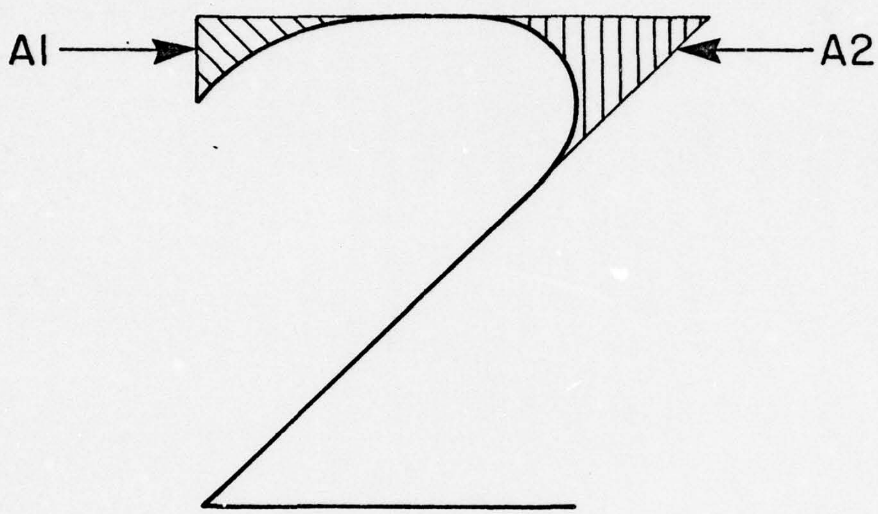


Fig. 4.1 Illustration of hypothesis 2

4.3 EXPERIMENT 1 - LABELING

The Versatec version of the characters in Fig. 3.6 were affixed to posterboard, one row of characters per board. The characters were 1.3 inches high and 1 inch wide with spacing of 2 1/2 inches between each and 1.85 inches margin above and below. Ss (subjects) were not shown the range of characters they would see, but were told they would see 2s and Zs. All Ss in this and subsequent experiments were recruited from the M.I.T. community and were paid a candy bar for participating in the experiment. S was seated at a desk and was given a short motivational explanation of the experiment. Each S was told to label each character from left to right as a 2 or a Z, one row at a time, and then to do each row again labeling each character from right to left. Rows were presented to 24 Ss, a unique row order to each S. The second pass through the rows was in the same order as the first for each S. As S finished looking at each row, it was placed face down on the desk. E (experimenter) held the row of characters approximately 2 1/2 feet from S's eyes and recorded S's responses on a pad not visible to S.

The labeling probabilities P(Z) and P(2), (e.g. P(Z) is the probability of a character being called Z, estimated by dividing the number of times it was called Z by the number of times it was presented) were estimated for each character. The results are shown in Fig. 4.2.

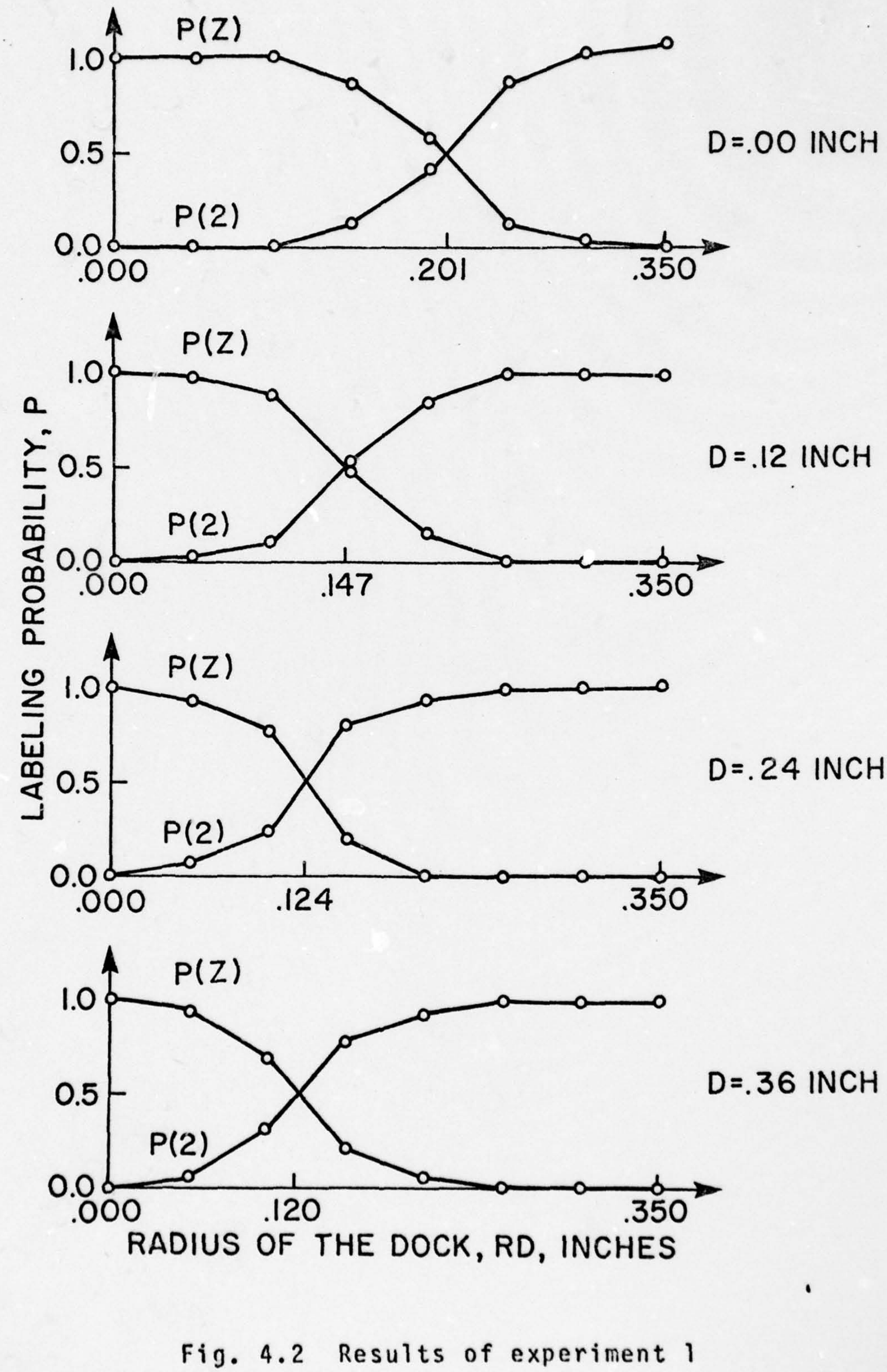


Fig. 4.2 Results of experiment 1

The boundary is estimated at the point where P(Z) = P(2) = .5 by linear interpolation of RD for each value of D.

Of the 32 characters presented, 18 were most often called 2, and 14 were most often called Z. This result is in agreement with the desired goal of presenting approximately equal numbers of 2s and Zs.

The following scheme was devised to evaluate hypothesis 1 and hypothesis 2: The experimentally determined boundary is plotted as a function of D and RD. The boundaries predicted by each hypothesis (a value of RD for each value of D) are calculated and plotted on the same coordinates, such that they pass through the two end-points of the experimentally derived boundary. The degree to which the hypothesized boundaries approximate the experimentally derived boundaries is then observed.

A plot of the hypothesis 1 boundary is a straight line through the end points of the experimentally derived boundary. A plot of the hypothesis 2 boundary requires that a linear relation between areas A1 and A2 be determined at the end-points of the experimentally derived boundary and that at each intermediate value of D a value of RD be determined which satisfies the relationship between A1 and A2. This has been done and is shown in Fig. 4.3.

Fig. 4.3 shows that hypothesis 2 predicts almost exactly the same description of the shape of the 2-Z intercharacter

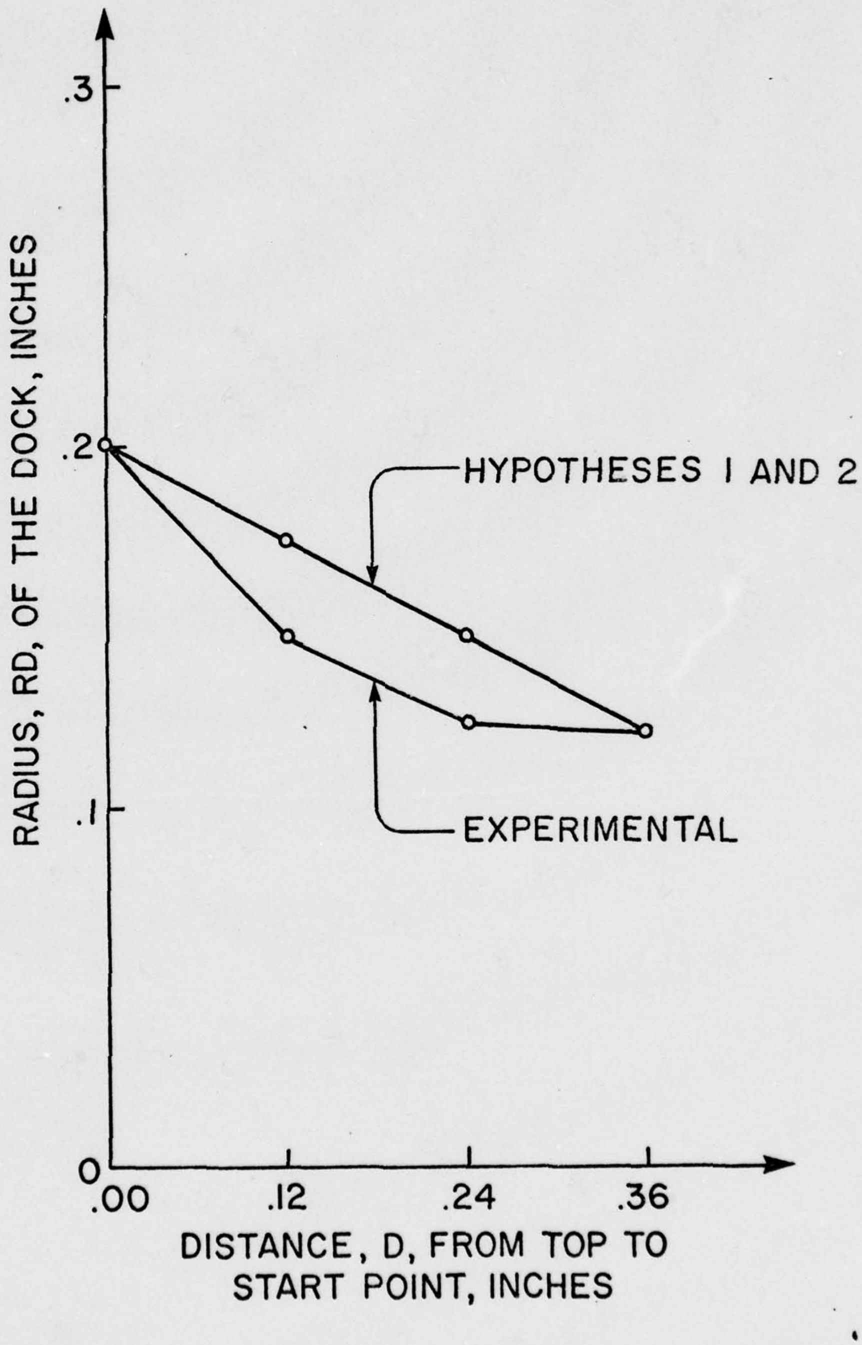


Fig. 4.3 Comparison of hypotheses for experiment 1

boundary as hypothesis 2; in the chosen parameter space the experimentally derived boundary is a concanve line, hypothsis 1 predicts a straight line and hypothesis 2 predicts an almost straight line. Experiment 1, however, was not designed to adapt the <u>S</u>s to the range of characters prior to experimentation; each row was presented as a unit, thus biasing the context for some rows [3,8]. This has the effect of shifting the boundary closer to the corner of some rows than they would have been if <u>S</u>s were completely adapted to the range of stimuli used. This suggests that the bend in the experimentally derived boundary may be due solely to the effect of biased context in the last two rows (D = .24 and .36 inch). At this time, therefore, neither hypothesis 1 nor 2 are rejected.

4.4 AN ADDITIONAL HYPOTHESIS

Riggs [6] performed color pattern after effect experiments with lines of varying degrees of curvature. Such experiments done with straight lines have previously pointed towards the existence of orientation sensitive regions of the visual system. Riggs experiments indicated that the orientation sensitive regions could not account for the after effects his subjects ovserved and that separate curvature sensitive regions probably exist. He suggests that specialized cortical cells exist for detecting the degree and direction of curvature.

Based on psychological evidence that the visual system is directly stimulated by curvature, it is desirable to formulate the third hypothesis in terms of curvatures. Two measures of curvature are commonly employed which I have called curvature and roundness. Curvature is zero for a straight line and infinite at a discontinuity in slope. Roundness is the radius of curvature and is infinite for a straight line and is zero at a discontinuity in slope. Therefore they are inverse measures; if R = roundness then 1/R = curvature. Now a good Z has zero curvature of the top left and zero roundness of the top right. As it becomes a good 2 the curvature of the top left, 1/RS, increases to some finite maximum as does the roundness of the top right, RD. It appears that the boundary may be a linear function of curvature and roundness, therefore hypothesis 3 is formulated; that the boundary is a linear function of 1/RS and RD.

After formulating hypothesis 3, experiment 2 is conducted.

4.5 EXPERIMENT 2 - REACTION TIME AND LABELING

Experiment 2 gathered both labeling and reaction time data. The characters of Fig. 3.6, 1.3 by 1 inch, were each mounted on a card, 4 inches high and 6 inches wide. Ss were shown a range card like Fig. 3.6 (without the labels) as they received their instructions. Ss were told they would see each character on the range card one at a time in random order and

they would have to respond to each character by calling it a 2 or a Z. They were told that they would see each character for 1/2 second in the tachistiscope upon depressing and releasing a button. They were told that their voice would be recorded in order to verify their responses but not that their reaction times would be recorded. Ss were told that this was not a test of ability and that there are no right or wrong answers. Fourteen Ss were tested and were paid a candy bar for their participation. Ss were shown a short series of cards to acquaint them to the procedures. When they felt comfortable with the equipment and procedure their experimental session began.

Ss viewed a blank prestimulus field of equal background intensity when not viewing a character in the tachistiscope. Stimuli were placed in random order according to a table of random numbers and presented one at a time. A digital timer displayed to E the time from onset of character viewing to onset of verbal response. E recorded stimulus number, label and reaction time for each character.

Labeling probabilities P(Z) and P(2) were determined for each character presented. The results are shown in Fig. 4.4 for each value of D. Intercharacter boundaries are estimated at the point where P(Z) = P(2) = .5 by linear interpolation.

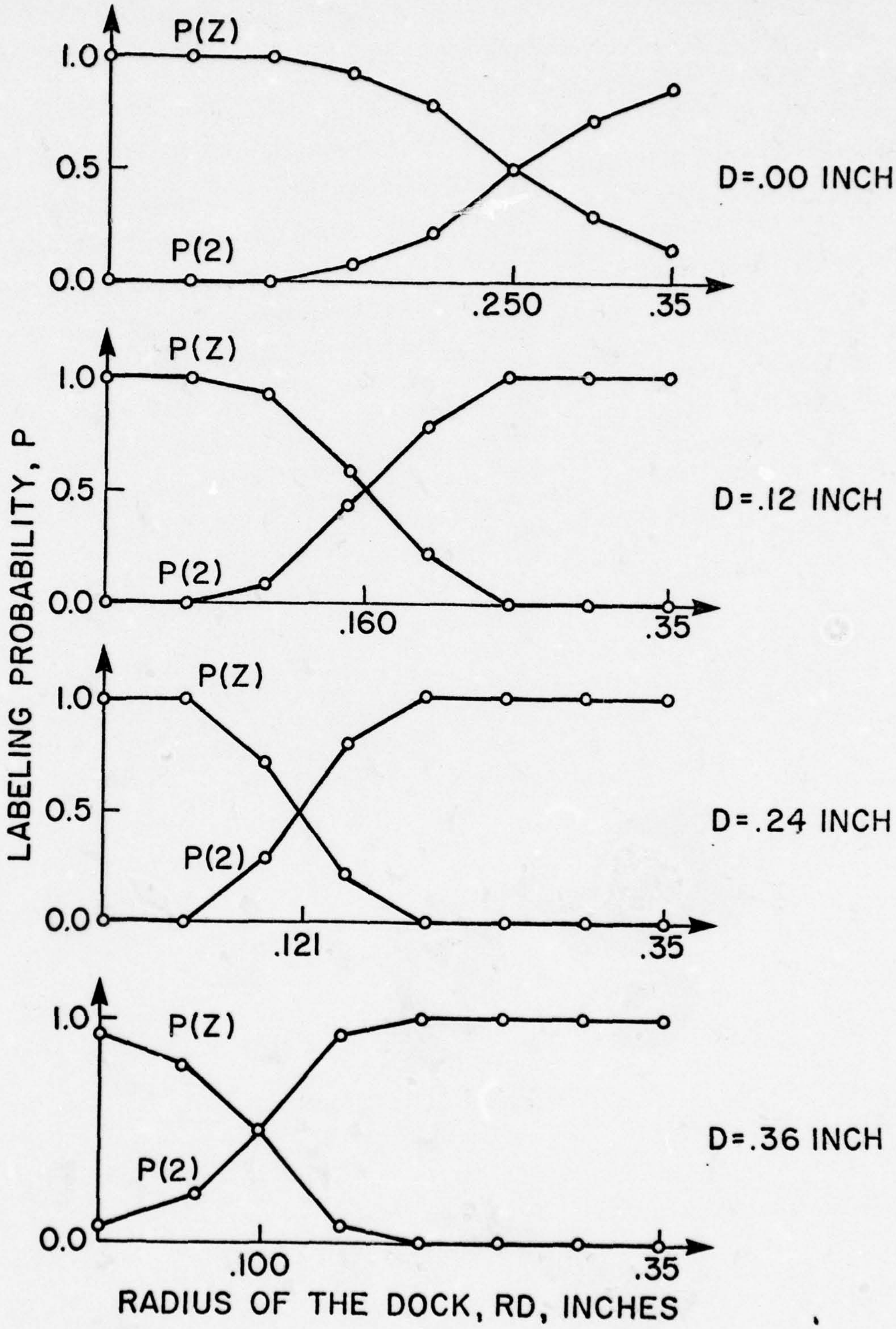


Fig. 4.4 Labeling probability curves, experiment 2

Hypothesis 1, 2 and 3 are now evaluated as in experiment 1 and are plotted in Fig. 4.5. It can be seen that the general shape of the experimentally derived boundary remained the same as it was in experiment 1, i.e., concave. The predicted shift of boundary occurred downward for the value D equal to .36 inch, however, the boundary shifted insignificantly for D equal to .24 inch and shifted up for D equal to .12 and 0.0 inch. Although the boundary did not straighten out, it is not unreasonable that it did not; the row with D equal to 0.0 inch was biased by context in experiment 1 with the presence of more Zs than 2s just as the row with D equal to .36 inch was biased by the presence of excess 2s.

Neutrality of range was maintained; of the 32 characters presented the consensus of Ss was that 14 were Zs, 17 were 2s and 1 was about equally likely to be either.

It is apparent from Fig. 4.5, that hypothesis 3 is far superior to either hypothesis 1 or 2 in predicting the correct shape of the boundary. This supports the contention that the curvature of the shores of an INLET are directly related to its functional SEGMENTATION.

The average reaction time for each character was calculated and is shown graphically in Fig. 4.6. The characters in the rows D equal to .24 and .36 inch have clearly defined maximum values of reation time for RD equal to .10 inch

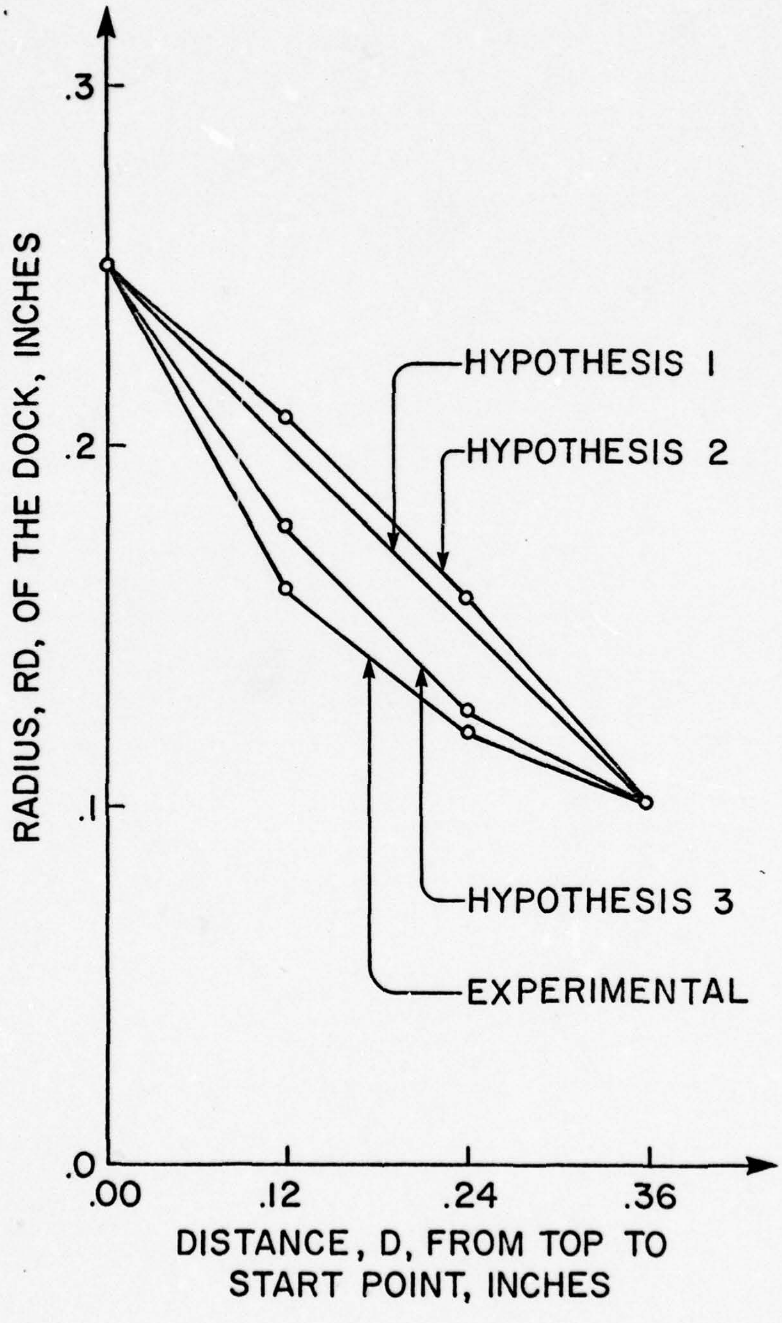


Fig. 4.5 Comparison of hypotheses for experiment 2 - labeling

which is in close agreement with the experimentally determined labeling boundary. The reaction time curve for D equal to .12 inch has a primary and secondary maximum value, the former being in close agreement with labeling data. The curve for D equal to 0.0 inch has a maximum value at RD equal to .30 inch which is not in close agreement with previous data. Furthermore, the shape of the curve indicates that reaction time is maximized over a region from RD equal .20 to .30 inch. To check labeling boundaries for each subject each row are determined and then arranged. The results of this calculation are shown in Table 4.1.

VALUE OF D	CORRESPONDING VALUE OF RD
0.00	.261
0.12	.154
0.24	.100
0.36	.100

Table 4.1 Experiment 2 - reaction time boundaries averaged over subjects

In light of the shape of the reaction curve in Fig. 4.7, the value of RD equal to .261 seems to be a more reasonable estimate than .30. The other values of RD are in close agreement with those of Fig. 4.6.

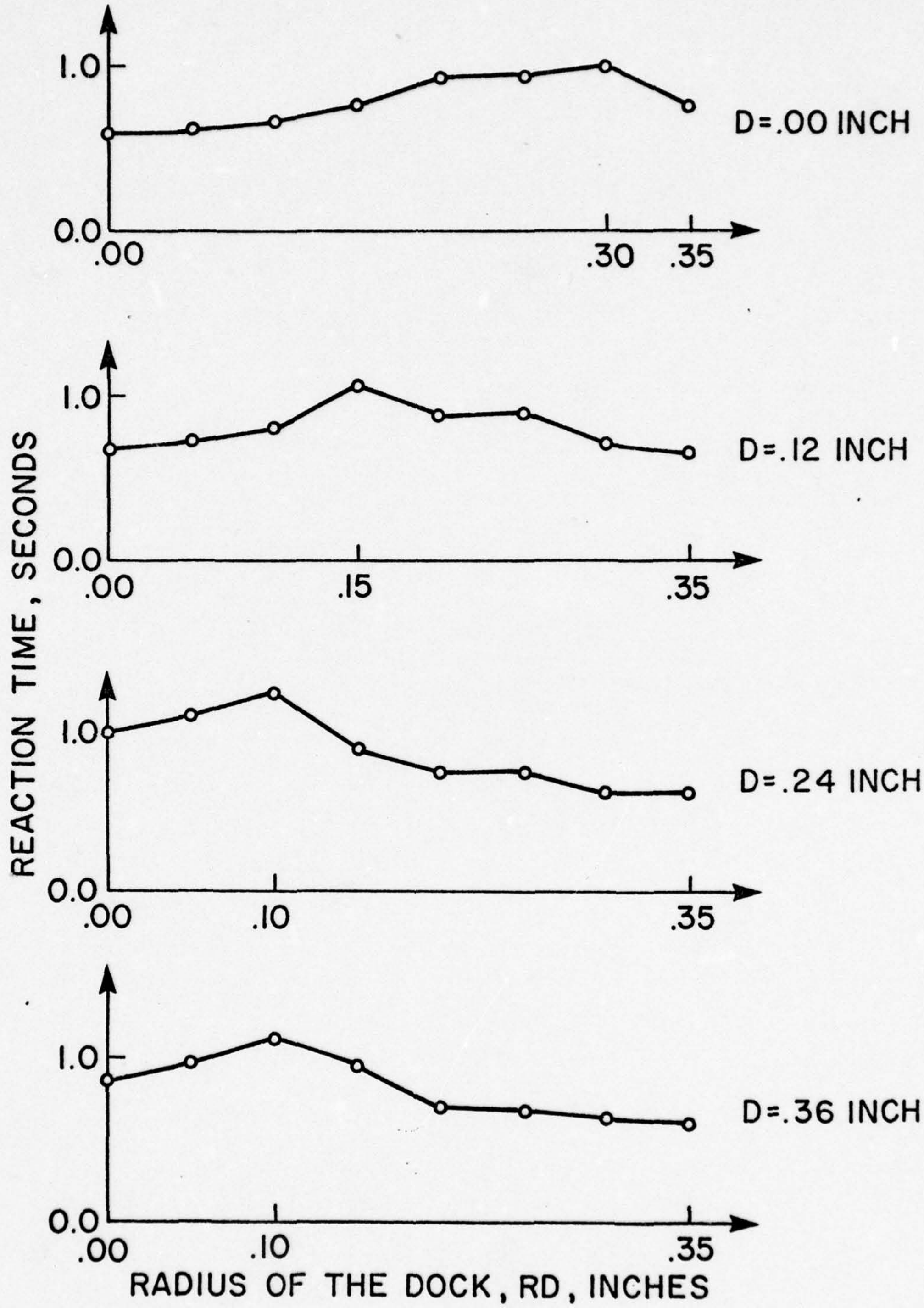


Fig. 4.6 Experiment 2 - reaction time averaged over characters

Deriving information from each subjects data and then averaging that information over the subjects is a standard technique for making significance tests between sets of data. Table 4.2 shows that averaging labeling boundaries over subjects yields estimates of RD consistent with data averaged over characters.

VALUE OF D	CORRESPONDING VALUE OF RD
0.00	.255
0.12	.161
0.24	.121
0.36	.089

Table 4.2 Experiment 2 - labeling boundaries averaged over subjects

A two-tailed dependent t-test indicates that there is no significant difference between the estimates of RD in Tables 4.1 and 4.2 at the .05 level.

The boundary determined from the reaction time data averaged over characters is compared in Fig. 4.7 with the boundaries that would be predicted by the relationships of hypotheses 1, 2 and 3. The hypothesis 3 boundary again gives the closest approximation to the shape of the experimentally determined boundary.

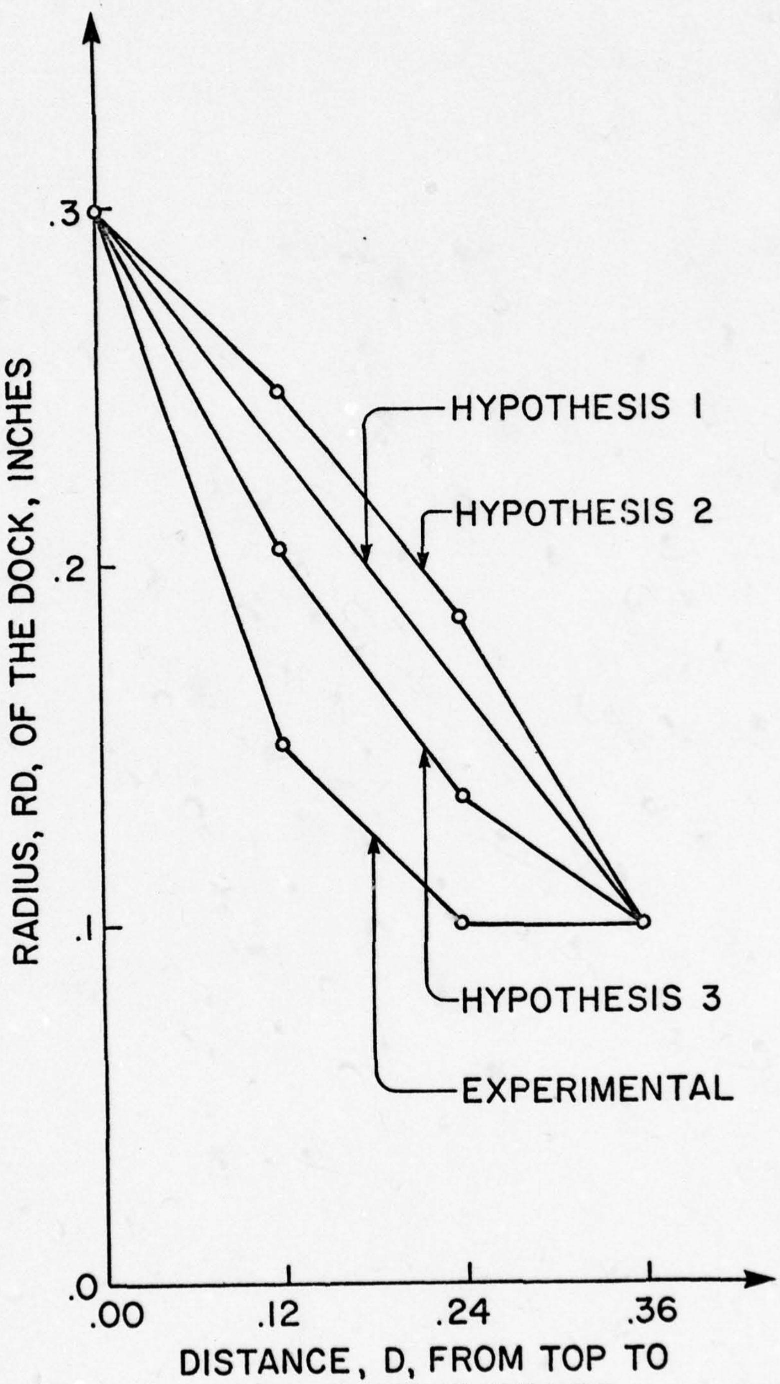


Fig. 4.7 Comparison of hypotheses for experiment 2 - reaction time averaged over characters

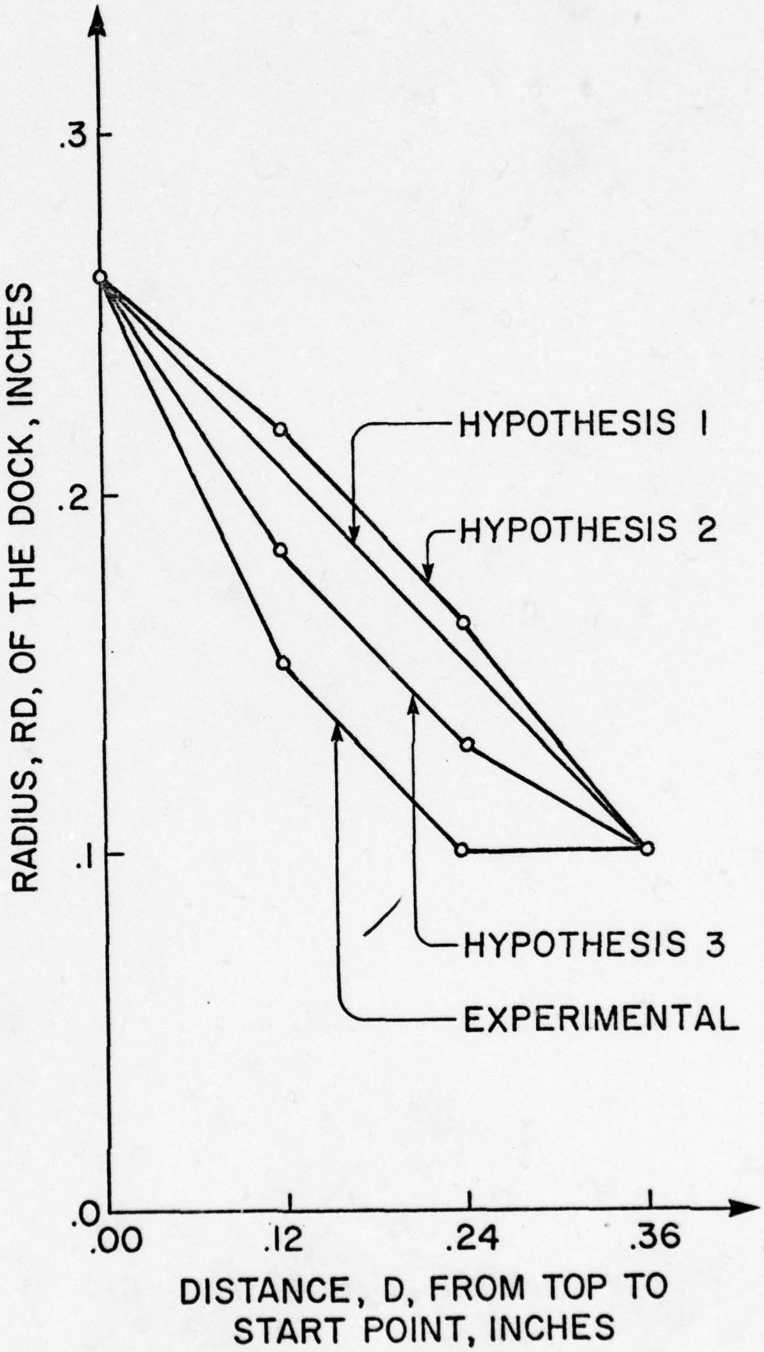


Fig. 4.8 Comparison of hypotheses for experiment 2 - labeling derived from reaction times averaged over subjects

The boundary determined by averaging boundaries derived from reaction times over subjects is plotted in Fig. 4.7 . Again the 3 hypotheses are compared and it is seen that hypothesis 3 provides the best approximation to the shape of the experimentally determined boundary.

4.6 EXPERIMENT 3 - GOODNESS

It is expected that the results of this experiment will not differ significantly from those of experiment 2 and that hypothesis 3 will be further supported.

The same characters used in experiment 2 were used in experiment 3. Ss were shown a range adapting card as in experiment 2 while receiving their instructions. Ss were told they would see each character on the range sheet one at a time in random order and that they should respond by rating each character as to how well it represents a Z (2 for half the Ss). When completed the Ss were told to view each character again rating how well they represent 2s (Zs for the same half of the Ss). Ss were told to hold each card while rating the character upon it. Ss were told there was no time limit for rating the character, that there are no right or wrong answers and that this is not a test of ability. Ss had little difficulty understanding or following the instructions. Cards were presented in random order to each subject for their rating as Z (or 2). Cards were presented in the same order to each subject for their rating as 2

(or Z). E recorded the sequence of presentation and rating of each. There were 22 Ss tested. The mean rating for each character as a Z and each character as a 2 was calculated. The results are shown in Fig. 4.9. Estimated boundaries occur at goodness curve crossings and are calculated by linear interpolation between points. Another estimate of the boundary is obtained by deriving the intercharacter boundary (called derived labeling) for each subject and averaging these points over subjects. This has been done and is compared in Table 4.3 with the previous estimates.

D	RD (by averaging over characters)	RD (by averaging over subjects)
0.00	.269	.262
0.12	.201	.209
0.24	.146	.142
0.36	.127	.133

Table 4.3 Experiment 3 - boundaries determined from goodness ratings

The data averaged over subjects are used in comparing the results of experiment 2 and 3. A two-tailed independent t-test on the data indicates there is no significant difference between them at the .001 level. The agreement between data averaged over subjects and over characters in both experiments and the statistical agreement between data averaged over subjects suggest that there is no significant difference between the boundaries of experiment 2 (labeling averaged over characters) and experiment 3 (goodness averaged over characters).

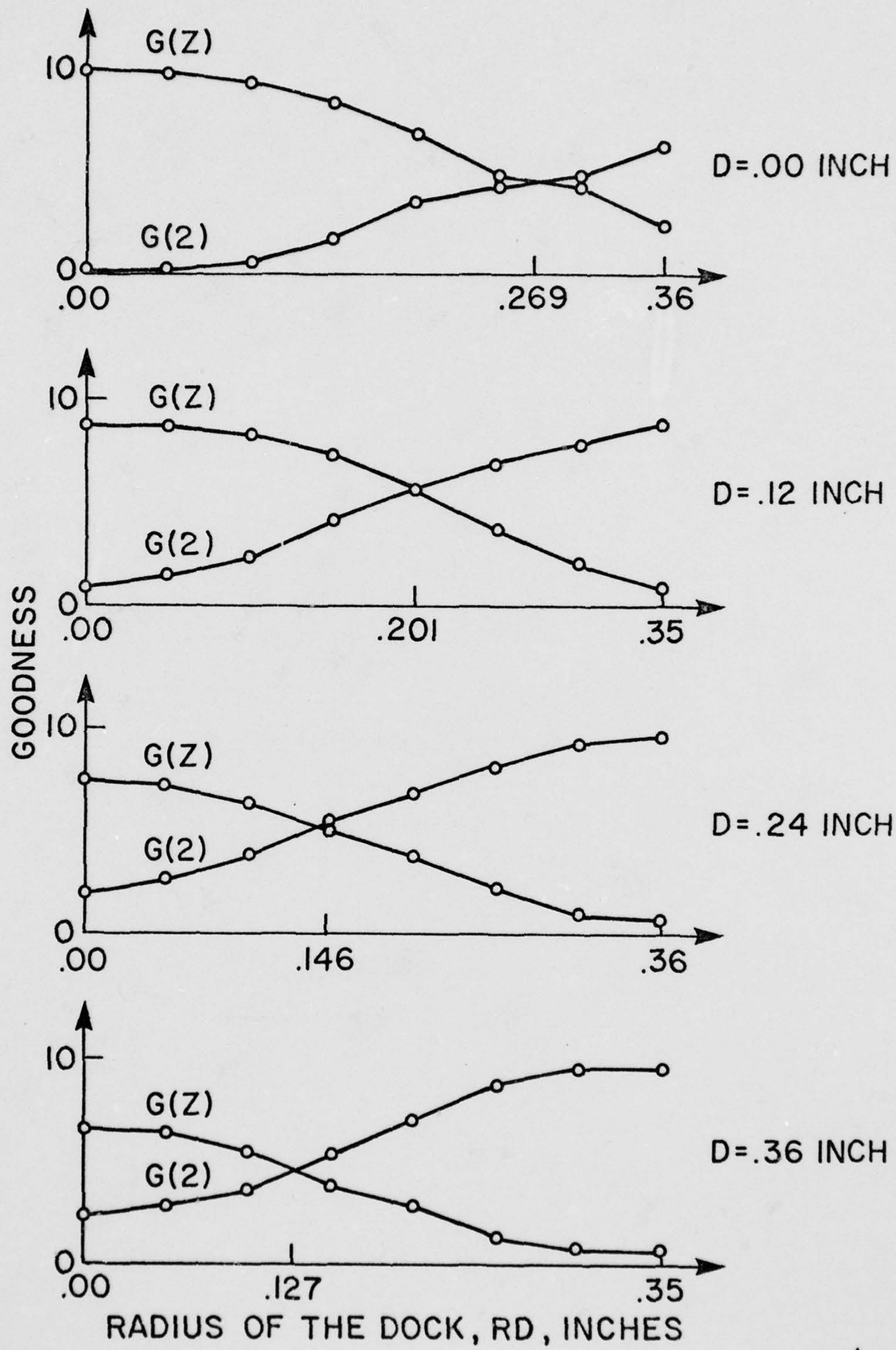


Fig. 4.9 Experiment 3 - goodness

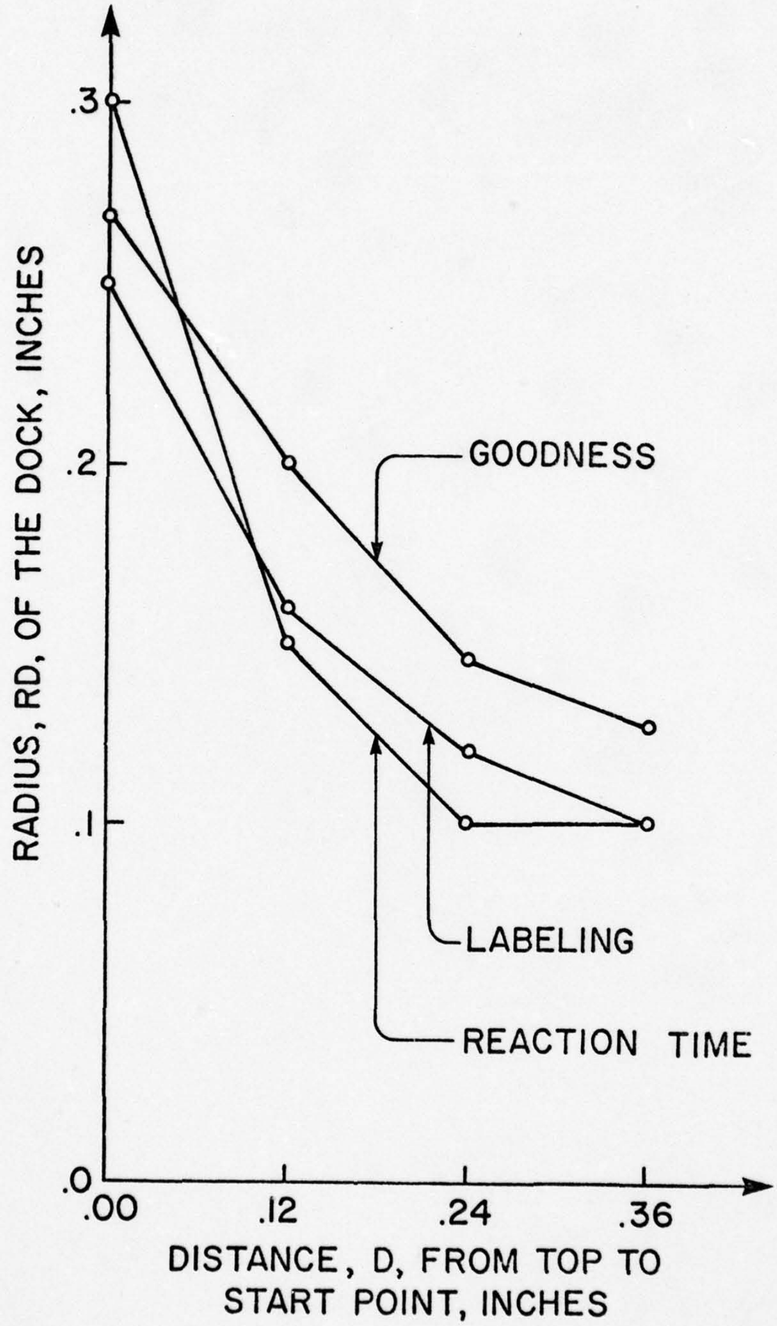


Fig. 4.10 Comparison of boundaries from experiments 2 and 3

This is further supported by comparing the graphs of the goodness, labeling and reaction time boundaries (each determined by averaging over characters) as shown in Fig. 4.10. Each of tbe three boundaries convey essentially the same information:

1. Character label depends primarily on the shape of its upper dock.

2. A little bit of curvature of the upper shore has a great effect on character label and

3. Increasing amounts of upper shore curvature have a decreasing effect on character label.

At most, the boundaries of Fig. 4.10 differ by one character which, from Fig. 3.6, appears to be a barely perceptible difference.

Finally, the goodness data, averaged over characters is plotted and compared with the boundaries of hypotheses 1, 2 and 3 that pass through the two end points. This is shown in Fig. 4.11.

The hypothesis 3 boundary provides an excellent approximation to the shape of the experimentally determined boundary while the other two are not nearly as close.

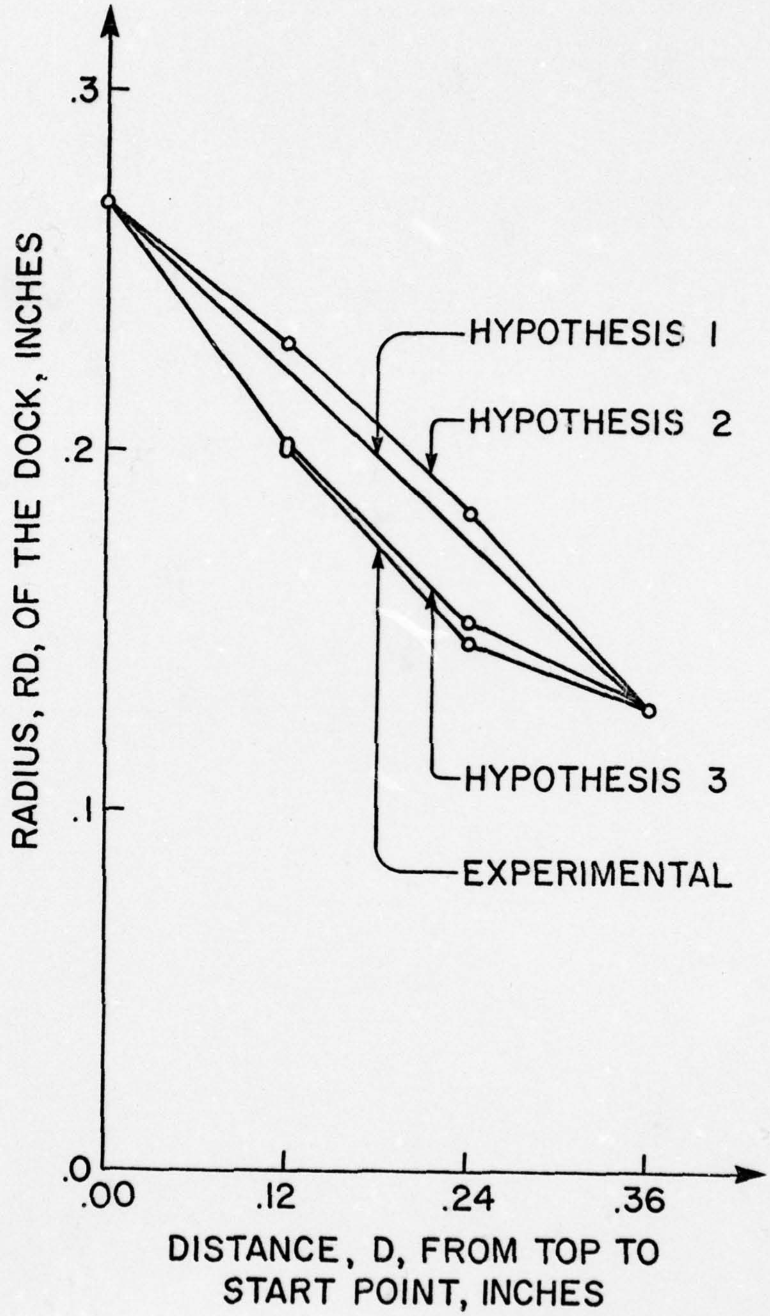


Fig. 4.11 Comparison of hypotheses for experiment 3 - goodness averaged over characters

4.7 DISCUSSION

Experiment 1 indicated the general location and shape of the intercharacter boundary in the chosen feature space between 2 and Z. Subsequent experiments showed that the range of characters in some of the rows presented as a unit in experiment 1 were biased and caused an undesirable shift in the boundary. Subsequent experiments were consistent with each other. Subsequent experiments were consistent with experiment 1 in that the shape of the boundary in the chosen feature space is concave. Neither hypothesis 1 nor 2 predict a concave boundary. Hypothesis 3 does predict a concave boundary which is very close in shape to the experimentally determined boundaries.

The results of experiments 2 and 3 may be combined to produce a single estimate of the intercharacter boundary. The derived labeling obtained from individual subject's reaction times, direct labeling and goodness ratings are averaged over subjects. Since two results per subject are available from experiment 1 and one each from experiment 3, the results of experiment 3 are weighted to effectively maintain one result per subject. These final results are shown in Table 4.4 and are the description of the experimentally determined intercharacter boundary between 2 and Z.

D	RS (from D)	RD
.00	∞	.266
.12	1.107	.183
.24	.641	.130
.36	.527	.118

Table 4.4 Final description of the intercharacter boundary from experimental data in inches

From the first and last values of Table 4.4 and hypothesis 3, a PFR for SEGMENTATION may be tentatively described as follows:

$$\left[\text{Functional SEGMENTATION: } \frac{.078}{RS} + RD \underset{\text{present}}{\overset{\substack{\text{not}\\\text{present}}}{\gtrless}} .266 \right]$$

In Table 4.5 the PFR is compared with the experimental results which were averaged over characters. The information of Table 4.5 is plotted in Fig. 4.12.

	Values of RD for D equal to:			
	.00	.12	.24	.36
	.201	.174	.146	.120
Labeling	.250	.160	.121	.100
Reaction time	.300	.150	.100	.100
Goodness	.269	.201	.146	.127
PFR	.265	.195	.144	.118

Table 4.5 Comparison of the SEGMENTATION PFR with experimental results averaged over characters

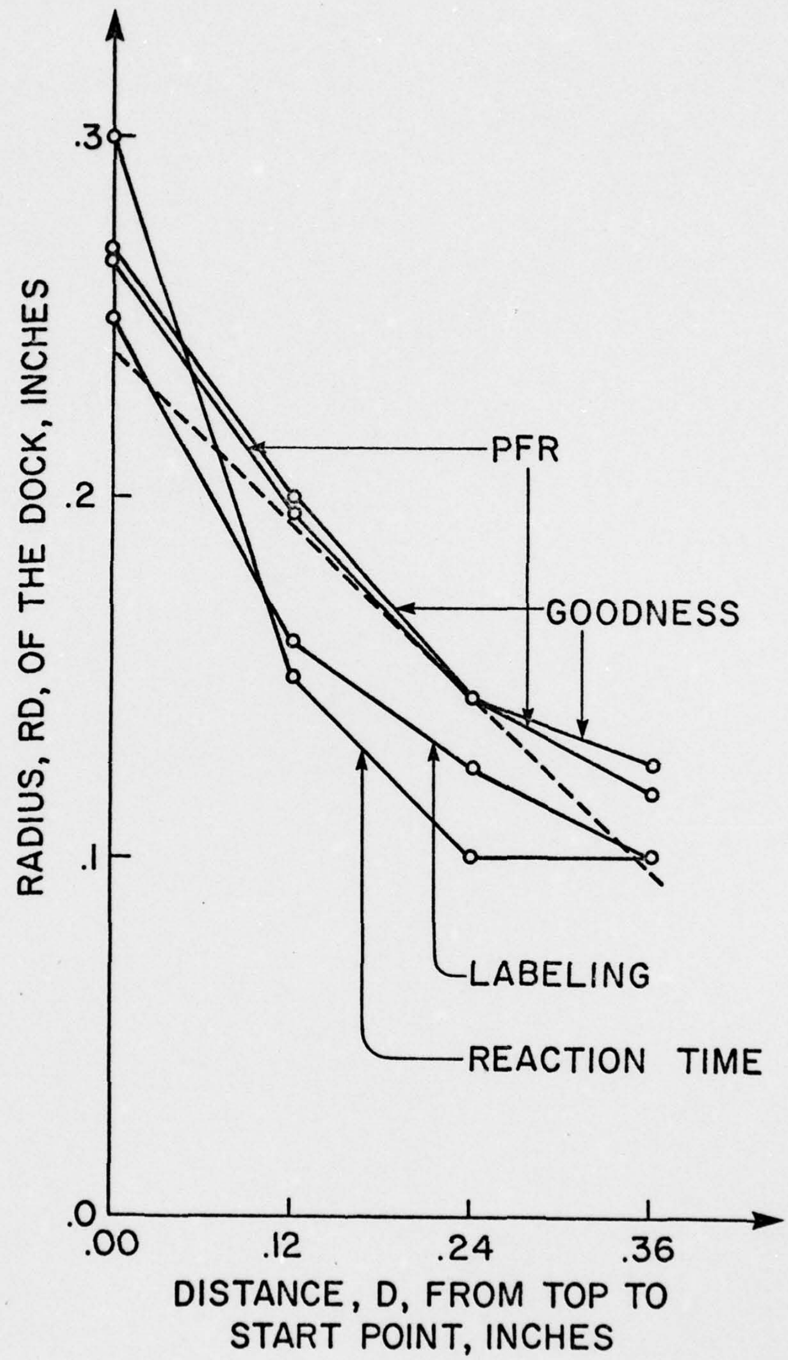


Fig. 4.12 Comparison of the SEGMENTATION PFR with experimental results averaged over characters

A PFR in the form suggested by either hypothesis 1 or 2 predicts a boundary which lies outside the range of experimental results for at least one value of D as shown by the dashed line in Fig. 4.12. Therefore, the PFR, in its present form, is the best choice of the forms considered.

CHAPTER 5

SUMMARY AND CONCLUSIONS

5.1 SUMMARY AND CONCLUSIONS

The developing theory of character recognition has been extended beyond the scope of letters to include Arabic numerals as well since letter-numeral pairs form a significant class of difficult pairwise cases. The method used to determine the underlying identity of characters is one of intercharacter boundary exploration through the use of ambiguous characters. Functional attributes and rules which map physical attributes into functional attributes are proposed and determined empirically through the use of psychophysical testing. A rule is initially determined in neutral context. Graphical context is then used to manipulate the rule, determine its effect and test the commonality of the attribute and rule across pairs.

In this paper the character pair 2-Z was explored. MARKERs such as crossbars, loops, spurs and curved bases were argued to be embellishments which are not included in the essential difference between 2s and Zs. It was argued that the essential difference between all 2s and all Zs is the functional SEGMENTATION of the upper INLET A character meeting the specifications other tha SEGMENTATION of the upper INLET as described in Section 2.5, was considered a 2 or a Z. This character, then, is a 2

if the upper INLET lacks functional SEGMENTATION and a Z if the upper INLET does have functional SEGMENTATION.

Possible physical variations were considered that could change a good Z to a good 2. The least physical change deemed necessary was considered to be rounding of the upper dock and some downward curving of the upper shore.

A two dimensional trajectory of characters was developed by computer to simulate these physical changes. These characters were used as stimuli in three psychophysical experiments.

Although experiment 1 had the effect of presenting several stimuli in biased context it did establish the basic nature of the boundary between plain 2s and Zs. Subsequent experiments substantiated the shape of the boundary. Experiments 2 and 3 were in good agreement with each other and provided the detail needed to state the PFR for SEGMENTATION.

From the results it is concluded that functional SEGMENTATION of the upper inlet distinguishes all 2s from all Zs. The PFR for SEGMENTATION, when it is fully developed, will probably involve the curvature of the shores and roundness of the dock.

The following PFR is consistent with the stimuli and similar characterforms with a height to width ratio of 1.3:

$$\left[\text{Functional SEGMENTATION: } \frac{.078W}{RS} + \frac{RD}{W} \;\underset{\text{present}}{\overset{\text{not present}}{\lessgtr}}\; .260 \right]$$

where RS is the radius of curvature of the upper shore, RD is the radius of curvature of the upper dock and W is the character width.

The PFR stated above is not claimed to be true in the general case. Curvature of the shore does not necessarily begin at its last half as with the experimental stimuli and radius of curvature generally is not constant along a curved line. It is expected that the stated PFR will be a specific case of the final PFR. It is concluded that SEGMENTATION is a difficult functional attribute to explore, but that roundness of the dock and curvature of the upper shore both contribute to its PFR.

This thesis followed a course of analysis which is recommended for initial investigation of other functional attributes. It is expected that, in solving the most difficult cases first, the subsequent cases will be trivial by comparison.

5.2 FURTHER RESEARCH

The usefulness of a PFR is a function of its simplicity in describing the essential distinguishing attribute of all characterforms in two classes and its applicability to more than a single pair of characters.

The first goal then should be to develop the SEGMENTATION PFR in a simple form, such that it distinguishes all 2s from all Zs as well as humans. This is no trivial task,

as an infinite number of physical variables potentially have an effect. The second goal, which should be incorporated with the first, is, while keeping the form simple to make the rule applicable to such pairs as 5-S and U-V. Finally, the ultimate test is to achieve similar changes in the PFR for the different character pairs for similar changes in context.

The computer character generator package is a useful tool in producing stimuli for psychophysical experiments. It has the advantage that the characters it generates have precisely controlled parameters. To be more useful the package needs some more capabilities. A modification, allowing the generated characters to be displayed on the refresh scope instead of the storage scope would make the MANUAL package equally applicable to both PDP-9 computers in the laboratory. Additional commands for segments such as polynomial curves or two arcs of different radii would be useful.

Ultimately, a set of PFRs must be found that describe a corresponding set of functional attributes which, taken together, uniquely specify the set of characters under consideration.

APPENDIX

This appendix summarizes the twelve physical attributes which correspond to the twelve functional attributes proposed by Shillman [8].

1. <u>SHAFT</u>

   A SHAFT is a vertical or horizontal line segment equal to the character height if vertical and equal to character width at that point if horizontal.

2. <u>LEG</u>

   A LEG is a line segment with one end attached to the lower half of a character and the other end free. It may be attached to the left, middle or right of the character. The free end may extend up, down or horizontally.

3. <u>ARM</u>

   An ARM is similar to a LEG except that it is attached to the mid or upper half of the character.

4. <u>BAY</u>

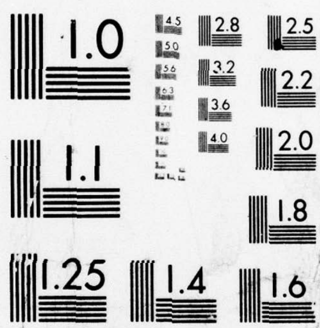
   A BAY is a concavity which has two "sides". The innermost region is termed the <u>dock</u> of the BAY and the "sides" are termed the <u>shores</u>. The ends of both <u>shores</u> (the ends furthest from the <u>dock</u>) are free. The BAY may be located in the top, bottom, left or right halves of the character or it may constitute the entire character. Its opening

AD-A054 755 MASSACHUSETTS INST OF TECH CAMBRIDGE DEPT OF ELECTRI--ETC F/G 6/4
SIMULATION METHOD OF FEATURE SELECTION FOR UNCONSTRAINED HANDPR--ETC(U)
MAY 77 R T BABCOCK

UNCLASSIFIED NL

2 OF 2
AD
A054755

END
DATE
FILMED
7-78
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

may be oriented up, down, right or left. It may consist of one, two or more segments. Two bays may occur in a character and be concatenated at their shores as in H or at their shores as in E.

5. CLOSURE

A CLOSURE is a completely surrounded region. It may be located in the bottom and/or top halves as in B or it may constitute the entire character as in D.

6. WELD

A WELD is formed by three line segments meeting at a point; two being smoothly connected and the third being "welded on". The WELD may be located at the top, bottom, left or right of the character. The third segment may extend down, up, left or right away from the WELD.

7. INLET

An INLET is similar to a BAY, except that it has only one free shore; the other shore is an extension of some other part of the character.

8. NOTCH

A NOTCH is a concavity which has no free shores. It may be located at the top, bottom, left or right of the character and always opens outward.

9. HOOK

A HOOK is a bend of approximately 180° in the end of a line segment as in J. The length of stroke

which is bent is usually less than one-half of the character's height (a bend greater than one-half of the character's height would probably be classified as a BAY).

10. CROSSING

A CROSSING consists of two strokes which intersect and cross as in X.

11. SYMMETRY

SYMMETRY regards curvature. SYMMETRY (of curvature) about a vertical axis distinguishes O and D.

12. MARKER

MARKER is the attribute which distinguishes O and Q.

BIBLIOGRAPHY

[1] Blesser, B., Shillman, R., Kuklinski, T., Cox, C., Eden, M., and Ventura, J., "A Theoretical Approach for Character Recognition Based on Phenomenological Attributes, "Int. J. Man-Machine Studies, Vol. 6, pp. 701-714, 1974. (Also presented at the First International Joint Conference on Pattern Recognition, Washington, 1973).

[2] Blesser, B., Kuklinski, T., and Shillman, R., "Empirical Tests for Feature Selection Based on a Theory of Character Recognition," Pattern Recognition, Vol. 8, pp. 77-85, 1976.

[3] Kuklinski, T., "Plastisity Effects of Handprinted Characters," Masters Thesis, M.I.T., 1975, unpublished.

[4] Naus, M., and Shillman, R., "Why a Y Is Not a V: A New Look at the Distinctive Features of Letters," J. Exptl. Psychol., Vol. 2, pp. 394-400, 1976.

[5] Neisser, U., and Weene, P., "A Note on Human Recognition of Hand-Printed Characters," Information and Control, Vol. 3, pp. 191-196, 1960.

[6] Riggs, L., "Curvature as a Feature of Pattern Recognition," Science, Vol. 181, pp. 1070-1072, 1973.

[7] Shillman, R., and Blesser, B., "Use of Ambiguous Characters in Measuring Functional Invariants," M.I.T. Research Laboratory of Electronics Quarterly Progress Report, Vol. 109, pp. 155-158, 1973.

[8] Shillman, R., "Character Recognition Based on Phenomenological Attributes: Theory and Methods," Ph.D. Thesis, M.I.T., 1974, unpublished.

[9] Shillman, R., Kuklinski, T., and Blesser, B., "Experimental Methodologies for Character Recognition Based on Phenomenological Attributes," Proceedings of the Second International Joint Conference on Pattern Recognition, Copenhagen, Denmark, IEEE Catalog Number 74CH0885-4C, pp. 195-202, 1974.

[10] Shillman, R., and Naus, G., "The Distinctive Features of the Letters O and D," M.I.T. Research Laboratory of Electronics Quarterly Progress Report, Vol. 118, pp. 233-238, 1976.

[11] Shillman, R., Kuklinski, T., and Blesser, B., "Psychophysical Techniques for Investigating the Distinctive Features of Letters," Int. J. Man-Machine Studies, Vol. 8, pp. 195-205, 1976.

[12] Suen, C., and Shillman, R., "Low Error Rate Optical Character Recognition of Unconstrained Handprinted Letters Based on a Model of Human Perception," to appear in IEEE Trans. on Systems, Man and Cybernetics, pp. 73-75, June, 1977.

[13] Wright, G.G. Neill, The Writing of Arabic Numerals, Scottish Council for Research in Education Series, No. 33, London: University of London Press, 1952.

IED

-78